重庆大学经济与工商管理学院研究生学位课程

《高等经济学》讲义

博弈论与经济模型

蒲 勇 健

二OO四年五月

原创作品 接受指正 未经许可 不得翻印

VISIT…

LANZAROTE Caliente.com

第二章　博弈与均衡

据《辞海》（夏征农，1999）所言，“博”有多种含义，通常所表达的含义是“大”、“广”、“通”，如“博学多闻”、“博古通今”；除此之外，还有“众多”、“丰实”之意；作为动词，其意为“换取，取得”，如“以博欢心，聊博一笑”等；但在古文中，“博”的本意又指“局戏”，即用六支筷子和十二个棋子进行赌输赢的一种游戏，谓之“六箸十二棋”。《论语·阳货》中有“不有博弈者乎？”，其中“博”乃指“六箸十二棋”。于是，“博”作为动词还隐含有“下棋”的意思。又据《辞海》本意，“弈”即“围棋”，所以，仅从字面上看，“博弈”即“下围棋”之意。在英文里，“博弈”的对应用词是“game”。据《英汉辞海》（王同亿，1982），“game”有“误乐或消遣”之意，也有“为达到一种目的所用的方法或策略”和“比赛，按规则进行的体力或智力竞赛”甚至“赌博”的含义。其中，“为达到一种目的所用的方法或策略”之意已经十分接近我们在“博弈论”中所使用的“博弈”二字的含义了。现在，我们使用词汇“博弈”或“game”早已不仅限于“下围棋”或“娱乐游戏”了，包括“下棋”、“玩牌”和其他许多智力游戏在内的对抗性游戏都可用“博弈”或“game”来指称（体力竞赛一般不用“博弈”指称了）。如果仅从字面上狭义理解，“博弈论”就是研究“下围棋”的理论；从较广义的含意上看，“博弈论”似乎指研究包括“下棋、玩牌在内的智力游戏”的理论，但实际上，我们现在使用“博弈”一词是在这样一种十分广泛的意义上加以理解的，即“博弈”是指“运用策略的各主体之间所进行的计谋互动过程”。这种定义既包括了“下棋、玩牌”在内的所有智力游戏，又包括了许多其他没有输赢的活动或过程，包括经济行为和军事对抗中的计谋活动。在正规的场合，通常按如下方式给“博弈论”下一个学术性定义：博弈论就是研究理性人行为选择的理论。这种定义出现在许多教科书中，已经成为了一种十分标准的定义。许多不同的学科都研究人的行为，如心理学、社会学，当然还有经济学，但我们今天所说的“博弈论”通常指经济学中的理性人行为选择理论。尽管如此，博弈论在军事学、生物学、系统科学等领域中也有十分广泛的应用，但从本质上说，“博弈论”是一种数学方法，而就数学来看，其应用对象的广泛性是十分自然的。与心理学、社会学中研究人的行为机制不同，博弈论研究的是“理性人”的行为。什么是理性人呢？按经济学家的定义，理性人就是追求效用最大化的个人。有人对理性人概念提出这样的批评：现实生活中的个人基于信息采集及处理的高成本和有限的计算能力，往往不可能在个人行为选择中实现其潜在的效用最大化，实际所完成的选择通常是次优（suboptimize）而非最优（optimize），因此，理性人概念是对现实中个人的非现实描写。但是，如果我们将个人在信息收集处理和计算过程中的困难理解为个人付出的成本，并且假设个人在决策时将会考虑到这种成本的存在且对行为选择的成本和效益进行权衡。那么，这种成本——效益分析的结果将决定个

人的行为选择。因此，在这种意义上，理性人假设并不与现实中的个人行为相背离，上面那种批评是缺乏逻辑基础的。当然，作为生物学意义上的个人是具有自由意志的人，他完全可以不按新古典经济学家所规定的效用最大化程序去进行个人行为选择，但这样一来，他可能尽做一些“蠢事”，不断地“吃亏”，最后被生活中的竞争所淘汰。如果一个企业家不追求效用最大化，他的企业就可能因为他所做出的一系列不明智的决策而出现经常性的亏损，最后也会在竞争中被淘汰出局。经济学家并不否认现实中的个人存在非理性行为或自由意志，但经济学家只是将研究的注意力集中在个人的理性行为方面，也就是说，经济学家只研究理性人的行为或人的理性行为，因为只有作出这种限定之后，经济学研究才有可能获得明确的结果和对经济现象作出预测，而预测能力是科学理论应具有的根本功能（关于这一点，我们将在下面加以展开）。

博弈论是一门十分年轻的科学，其研究的起源可以回溯到第二次世界大战早期，当英国海军和德国潜艇玩猫捉老鼠游戏时，需要对游戏更深刻地理解才能更多地赢得胜利。他们发现并不是依靠领航员或船长凭直觉作出的判断而采取了正确的行动。在他们应用了后来被称为博弈论的概念之后，英国海军极大地提高了对德国潜艇的打击命中率。而对德国潜艇的胜利又促使他们把博弈论应用到许多其他的战斗中去。因此，博弈论是在生与死的环境下得到验证之后，才在今天以系统理论的面貌出现在书本上的。在这以后不久的 1944 年，由天才数学家冯 •诺依曼与经济学家摩根斯坦合著的《博弈论与经济行为》（Von Neumann and Morgenstern，1944）标志着博弈论经典科学体系的诞生，“博弈”一词从此也远远超越了下棋玩牌之类的智力游戏含义而泛指更为广义的计谋活动。这一辉煌的巨著，预示了 20 世纪最伟大的科学成就之一的到来。它引发了经济、政治、军事战略、法律、计算机科学、甚至生物进化论等领域大量的科学论文。在这些领域，博弈论都能阐释出主要的发现。但博弈论在其十分短暂的发展过程中却有着十分曲折的经历。在博弈论诞生伊始之时，整个社会乃至学术界都不太了解它，以至于二战后颠沛流离于澳大利亚的博弈论大师、匈牙利人海萨尼（Harsanyi）居然在大学图书馆中看见博弈论学术专著被与描写下棋玩牌等游戏类的书籍混杂摆放在一起。在这样一种不识博弈论为何物的环境里，海萨尼本人的命运可想而知，在失业和学术价值得不到承认的双重打击下，他辗转去了美国。1994 年海萨尼与纳什（Nash）和泽尔腾（Selten）分享当年的诺贝尔经济学奖，当时的美国报刊就以“澳大利亚是博弈论的最大输家”为题对海萨尼在澳大利亚的这段心酸事作了报道。澳大利亚因不识海萨尼所做工作的价值而失去了拥有一位诺贝尔奖得主的机会，损失当然巨大啦！

2.1 作为经济模型的博弈论

2.1.1 游戏博弈与经济互动的异同

诸如下棋、玩扑克牌在内的各种智力游戏都有一个共同特点，即参与游戏的各方之

间存在着输赢。在游戏进行之中，一方赢得的就恰好等于另一方输掉的。譬如，在国际象棋比赛中，一方吃掉对方的一个棋子，就意味着该方赢了一步而对方输掉一步，我们称这种博弈为“零和博弈”（zerosum game）。倘若我们在象棋比赛中作出这样的规定：当一方吃掉对方的一个棋子时，对方应输给该方一分钱，并用“支付”（Payoff）一词表示双方各自输赢的总和，则在比赛进行过程中以及比赛结束时双方的“支付”相加总等于零。所谓“零和博弈”的概念就是由此而来的。我们称博弈中的参与主体为“局中人”或“参与人”（player），而各局中人的所得（或所失）为各自的“支付”，则当且仅当所有局中人的支付之和在博弈进行过程中及博弈结束时恒为零时，博弈是“零和博弈”或“零和的”。显然，倘若我们在各局中人的支付中同时分别加上或减去相同的一个数量，博弈的过程特征则并没有什么不同，但这时各局中人支付之和并不等于零而恒等于一个常数，此时我们称博弈是“常数和博弈”或“常数和的”或“恒和博弈”或“恒和的”。从博弈的本质特征上看，常数和博弈与零和博弈并无什么不同。因此，我们习惯上将常数和博弈称为零和博弈。这是因为，效用函数在加上或减去一个常数后仍为同一偏好序的效用函数，而当常数和博弈中的各局中人支付加上或减去一个常数后，常数和博弈就变成了零和博弈（当然，零和博弈本身就是一种常数和博弈）。尽管下棋和玩牌等游戏博弈通常都是零和的（在玩成平局时，各局中人的支付都为零，这些支付的总和仍为零），但在经济活动中的许多互动却是非零和的。在现实的经济生活中，人们常常将因合作带来的额外收益称为“双赢”（win–win），即合作往往会给经济主体间带来 1 加 1 大于 2 的结果，有时人们也将这种合作行为称为“正和博弈”。相反，不合作行为往往带来 1 减 1 小于零的负效应。可称这种不合作行为为“负和博弈”。当然，这种用语仅仅是一种形容意义上的表称，与我们所说的“博弈”还不是一个概念。因为将“正和博弈”中的各个局中人支付减去一个足够大的数岂不就变成了“负和博弈”，而这又是与“正和博弈”和“负和博弈”概念的原意相左的。但是，经济互动中的许多行为过程并非导致一方所得就是另一方所失，有可能双方都因某种行为的选择而同时有所得或同时有所失，因而这些互动并非是“零和的”。这是棋牌类游戏与经济互动的区别。但是，与棋牌类游戏相类似的是，经济互动中的每一个行为都应被视为一种“理性的选择”，经济互动是理性人行为选择的互动（对“理性人假设”持有异议的读者可回过头去重新读读本章开始时的内容）。同时，无论是下棋玩牌还是经济互动，都存在着“信息”和“势”方面的问题。在国际象棋比赛中，倘若两个局中人都具有同样的理性程度，则数学家 Zermelo（Zermelo，1913）曾证明比赛的结果必为平局。但是，现实中的对弈结果并非总是平局，这是由于局中人的理性程度通常是不对称的，这种不对称或源于计算能力的不对称或归于信息掌握的不对称。譬如，在陆战棋对弈中，信息的获取与有效利用是一个十分重要的因素。当你的一个“师长”被对方“吃掉”后，如果吃掉你的“师长”的对方棋子并未同时被报销，则你由此可推断对方棋子至少是一个“军长”，即或者是“军

长”或者是“司令”，不会是别的。这样，尽管你的一个棋子被对方“干掉”，但你却同时获取了一个额外的信息：“干掉”你的“师长”的那个棋子必定是“军长”或“司令”（尽管你不能观察到那个棋子的真实“身份”）。又如在扑克牌游戏中，局中人不一定能记住已经打出的所有牌张，但局中人的记忆能力是不一样的，有的能记住更多的已出牌张，有的却差一些。这样，记住更多已打出牌张的局中人就比别的局中人拥有更多的信息，从而能打出更好的牌。另一方面，无论是下棋还是玩牌，不同局中人的理性程度都存在一定的差异，如计算能力上的差异等，这种差异将在很大程度上决定着最终的胜负。如在著名的春秋战国故事“田忌赛马”中，齐国大将田忌的理性程度低于齐威王，而齐威王的理性程度又低于孙膑，所以田忌输给了齐威王，齐威王又输给了经孙膑点拨之后的田忌。

“势”指博弈开始之前局中人拥有的“禀赋”，如在扑克牌游戏中当摸牌过程完成后的花色品种分布及有无王牌等。拥有好的“势”对博弈取胜至关重要。当你摸到了大小王牌后，你会喜形于色。因为取胜的把握增大了几分！中国古代著名军事理论家、“兵圣”孙武子在其《孙子兵法》中写道：激水之疾，至于漂石者，势也。……故善战人之势，如转圆石于千仞之山者，势也（黄朴民，1989）。正是由于“势”是如此重要，在战争开始之前，敌对双方都会花大力气进行“造势”，以形成有利于己方的形势，比如抢占至高点、鼓舞士气等。故孙子曰：是故善战者，其势险……故善战者，求之于势，不责于人，故能择人而任势。任势者，其战人也，如转木石……（黄朴民，1989）。在围棋博弈中，开始几招几乎都着眼于“造势”，抢占有利“地盘”，然后才有捉对“厮杀”。正是由于“势”这种不对称条件的存在，围棋博弈中“执黑先行”者通常就有“先动优势”；而在足球比赛中，开球时通过裁判抛硬币来决定哪一方开出第一球，这是为了公平地分配比赛双方的“势”。当然，在象棋和陆战棋博弈中，仅包括谁先走一步的“先动优势”，因为双方在其他条件方面均相同。

类似于下棋玩牌甚至战争中的计谋运用在经济活动中也随处可见。在商战博弈中，商家凭计谋取胜更被视为商战高招。孙子云：不战而屈人之兵，善之善者也（黄朴民，1989）。企业在竞争中也会运用“造势”来形成竞争优势。譬如，某些企业会抢先进入新兴行业并通过迅速扩大规模来阻吓竞争对手。因为其巨大的生产规模带来的规模经济优势使其拥有较大的价格下调空间，从而迫使潜在的竞争者不得不考虑价格战所可能造成的巨大损失，使其最终放弃进入行业的打算。

除此之外，信息不对称也会对经济互动产生重要的影响。

在发达国家的二手车（旧车）市场上（美国人俚称有问题的旧车为“柠檬”（lemon），故又称二手车市场为“柠檬市场”（the market for lemons）），一个常见的现象是二手车价格与新车价格相差很大，一辆仅使用过一年半载的旧车即使完全没有什么毛病，也只能在二手车市场上以比新车价格低很多的价格卖出去。这是由于，买车人对与旧车的质

量有关的信息掌握得很少，至少比卖车人掌握的有关信息少得多。无论旧车有无毛病，卖车人都会尽量夸大旧车的质量。倘若旧车确有质量问题，卖车人也会尽量隐瞒真情，而买车人也会事先悟出这番道理，因而不完全都相信卖车人所说的情况。这样，为了不致于吃亏上当，买车人总是会开出一个较低的价格，从而冲抵潜在的风险。所以，二手车市场的成交价格都较低。在许多场合，二手车市场甚至难以建立起来。倘若卖车人所欲出售的旧车的确是没有质量问题的好车，尽管卖车人说的全是大实话，但买车人怎么也不会完全相信他，买车人还是会开出一个较低的价格；于是，卖车人或者忍痛割爱只得以低价成交（可能他有十分紧迫的原因不得不将车卖掉），或者根本就不愿低价出售而退出二手车市场。相反，倘若卖车人欲售之车真的是存在严重质量问题，但他仍会王婆卖瓜自卖自夸，将破车说成是宝贝。尽管买车人不会完全相信他所说的，但仍然相信市场上也有一定的可能性存在好车，于是，他开出的买价可能高于车子的真实价值。而在这种价格上，卖车人会喜出望外，积极地向买车人推销，希望立即将破车卖出去。这样，市场上凡是十分乐意积极成交的卖车人所售之车很可能有问题，而买车人也会事先就明白这个道理。于是，买车人会改变主意，将开出的价格重新进一步向下调整。这种调整过程会不断进行下去，直到市场上的卖车人所兜售的全是破车为止。如果买车人不愿买一辆存在严重质量问题的车回家（可能基于安全考虑），则买车人一旦意识到市场上全是破车时，他们也会退出市场，其最终结果是市场消失。这种心理上的推理互动完全可以不经实际过程而完成，于是就没有人愿上二手车市场购车，因而二手车市场就难以建立起来。我们可以用一个数学例子来说明这个道理。假设开始市场上有一半的旧车是好车，另一半是存在质量问题的破车。假定好车价值 20 万元，破车价值为 10 万元，买车人对具体哪一辆车是属于好车（或破车）的信息是不完全的，但卖车人知道自己所卖之车到底是好车还是破车。买车人尽管不知道具体一辆车的质量情况，但他知道市场上仅有 50%的车是好车。于是，他开出的最高价为每辆 15 万元（(20+10)/2=15）。此时，卖好车的卖车人不愿成交，因为他的车价值 20 万元，以 15 万元成交是亏本生意，但卖破车的卖车人都会十分积极地成交，热情地向买车人推销，因为他的破车只值 10 万元，倘若可以按 15 万元成交，算下来会大赚一笔。假定买车人是理性的，他也明白这个道理，于是，他会将那些积极与他按每辆 15 万元成交的卖车人视为出售破车者，他因此会将市场上还在继续出售的车中存在破车的概率向上调整（他知道在每辆 15 万元价格水平上，好车已退出市场），直到将此概率上调到 1，此时，他认为市场上出售的全是破车（的确也如此）。这就如同历史上实行货币金银本位制时出现劣币驱逐良币的现象。如果买车人不想买一辆破车回家，则买者会退出市场，整个旧车市场消失。顺便提及，这种分析并非杜撰，而是存在真实的案例。经济学家阿克洛夫曾在一篇划时代的论文中给出了真实例子，也正是这篇论文给出的“二手车模型”大大提升了不完全信息博弈论在主流经济学中的地位（Akerlof, 1970）。正是买者与卖者对商品质量信息的不对称（卖

者知道的信息比买者多）造成了市场上劣质品驱逐真品的结果。用“二手车模型”可以解释大量政治、经济甚至社会现象。譬如，在政治场合，由于公众对官员们的清廉与腐败的属性具有不对称信息，可能造成贪官挤出清官的现象，从而导致政坛污浊、腐败成风（蒲勇健，2001）。在一个单位里，由于领导人对下属职员真实能力具有不完全信息，从而在待遇上亏待最有能力的职员而造成人才流失（正是由于这一原因，发达国家的一些企业实行“效率工资制”以避免这种“逆选择”结果的出现）。在金融市场上，金融机构由于对借款人的还款能力具有不完全信息，可能造成金融市场供求不均衡（近年来中国国有商业银行在实行资产负债比例管理后出现的“惜贷”现象即属此类）。对于谙熟博弈论的经济学家来说，这种可用“二手车”模型予以解释的经济及社会现象可谓俯首即拾。

正是游戏博弈与经济互动之间存在着那么多惊人的相似之处，启发经济学家开始思考如何运用游戏博弈的思想去刻划经济互动的过程与机制。游戏博弈与经济互动之间的相似点本质上表现为它们都是行为主体通过理性选择其行动或策略以达到预期效用最大化目标的互动过程，二者之间在这方面的一致性为经济学家开始运用博弈思维打下了基础。

2.1.2 科学理论与模型

经济现象十分复杂，经济学本质上是研究复杂现象的科学。这种复杂性表现在几个方面：首先，经济现象最终归结为个人的一系列行为选择，而即使我们暂且不考虑个人自由意志对行为选择随机性的影响，人与人之间的互动就是一个相当复杂的过程。类似于物理学中的多体问题，一个社会中众多个人之间的互动带来的相互作用结果往往是复杂的不断调整而不是一个终极的静态结果。我们知道，在物理学的多体问题中可能会出现混沌（chaos）现象，因此，一些非线性科学家试图运用混沌理论对诸如宏观经济波动和股票市场波动作出分析和预测（蒲勇健，1995；理查德・H・戴等，傅琳等译，1996）。其次，人的行为受到人的大脑意识的支配，而人的大脑意识及心理活动机理目前我们还未完全搞清楚，因而就难以对人的行为作出准确的预测。但另一方面，心理过程必定会通过其对人的行为选择的支配而影响到经济行为的最终结果（凯恩斯先生最早意识到人的心理过程对宏观经济影响的重要性）。这样，我们会对许多经济现象感到迷惑不解就是十分自然的事了。最后，许多经济活动表现出来的结果往往是人们在经济动机和人的其他社会性方面进行权衡或互动的综合效应，经济动机只是诸多形成因素中的一种，因而仅通过经济分析是难以给出完整解释或准确预测的。近年来，一些主流经济学家试图通过经济分析去解释或预测传统经济学没有涉入且是其他社会科学传统研究领域中的现象，并且这种努力还多少获得了一定成功（贝克尔，1997），但这种努力在其他社会科学界包括经济学界都招来了批评声，而这种行为本身又被斥之以“经济学帝国主义”

倾向，即一些经济学家将经济分析范围扩张到其他社会科学领域的倾向，被比喻为一个帝国入侵别国的行为。我们认为，这种批评有一定道理，试图通过经济学去完成对人的社会行为的完整描述可能是既不现实的又非必要的，或许是一种乌托邦式的幻想。如果是这样，经济学家就应该去做“经济学能够做的事”，而将其它留给其他社会科学家去处理。多样化不仅有利于社会科学的发展，而且也有利于经济学自身的发展。当经济学家只运用经济分析去剖析本质上综合了包括经济的、社会的、法律的以及其他诸多方面因素的社会经济现象时，他们所面对的巨大复杂性是可想而知的。

那么，经济学应如何有效同时又是有节制地对经济现象进行分析呢？从更为宽广的角度上看，科学家们应该如何去分析所观察到的现象呢？近代科学的发展在这种方法论意义上已有很大的突破，那就是通过“模型化”方法去理解、解释并预测事实现象。模型是对现实的抽象，它要展现出现实复杂系统的主要特点。生物学家用塑料人体模型来研究解剖学，这些模型有所有主要的器官——心脏、肺、肾等等。这种模型使生物学家能以一种简单的方式理解人体的这些重要器官是如何结合在一起的。当然，这些塑料模型并不是真实的人体，而且也没有一个人错误地把模型当成真人。这些模型是程式化的，并略去了许多细节。尽管它缺少了真实性——实际上正是由于缺乏真实性——但研究这些模型对了解人体如何活动是有帮助的。建筑工程师通过模型去考察各种风力和承载力对建筑物的影响，尽管建筑模型与真实的建筑物相比完全是不同的东西。

经济学家也用模型来了解世界，但不是塑料模型，而通常包括了图形和数学方程式。与生物学家的塑料模型一样，经济模型也忽略了许多细节，以便使我们能说明什么是真正重要的。正如生物学家的模型并不包括人体所有的肌肉和微血管一样，经济学家的模型也不包括经济的每一个特征。

从古代人的天圆地方学说到托勒密的宇宙模型到牛顿力学到现代量子论和相对论宇宙学都是使用模型来理解物理世界的例子。一个好的模型是在模型构造的现实性与简单性之间达到一种恰如其份的平衡。模型可以是“好”的或是“坏”的。假如抽象过分，模型就说明不了它想说明的现实系统；相反，如果抽象不足，这个模型就会太复杂，对洞察现实系统的运行就没有什么价值。模型愈复杂，它所容纳的因素就愈多，也就愈贴近现实，但付出的代价是研究者在运用模型进行预测或进行分析时会感到愈困难。相反，模型愈简单，运用模型进行预测或对理论进行验证时就愈省事，但可能出现谬误的可能性就愈大。科学大师们往往具有一种非凡的本事，他们能够在模型的现实性与简单性之间作出很好的权衡（trade–off），处理得恰如其份，从而能依据模型对事实作出准确的预测。

另外，模型愈复杂并不一定意味着它就愈接近实际情况。有时简单的模型往往更接近现实或更为正确。譬如，托勒密的由本轮和均轮构成的太阳系模型比哥白尼的日心说模型复杂得多，但事实证明后者才是正确的。爱因斯坦曾说：简单即美。在科学发展史

上的诸多例子说明简单的理论很可能就是更为正确的理论。这是因为，简单的模型往往抓住了事物的本质。在牛顿力学中，质点是一个不现实的假定，即没有体积但有质量的点，但它却是一个对事物十分成功的简化。在粒子物理的微观尺度和宇宙学宏观尺度上，我们人类的直接观察能力是十分有限的，但我们可以通过数学模型如量子理论、夸克模型、甚至超弦模型去理解微观世界的行为，也可以通过大爆炸宇宙模型去追溯宇宙的创生、演化和预测其未来的归宿。经济模型更是对经济行为的高度简化。由于经济现象的高度复杂性，经济模型必须对许多与所研究问题不太直接相关的细节进行简化，而这种简化是通过诸多假设来完成的。在经济学研究中，假设是十分重要的。不同的假设其实对应于不同的经济模型，不同的假设给出不同的预测结论。许多经济学家在经济问题上的分歧源于他们潜在地依赖于不同的假设。采用数学模型，我们可以清楚地看出假设与预测结论之间的对应关系。因此，经济学家们的争论通常可以回归到他们在相关假设上的不同，只要检验各自的假设与现实的一致性，我们就可以对经济学家之间的分歧作出评论并在不同的经济政策之间进行选择，这就是现代经济学普遍使用数学语言的原因。数学方法在科学研究中往往还会为科学家带来完全凭直觉无法获得的正确结果。譬如，爱因斯坦在完成了引力场方程构建之后，发现宇宙解是膨胀解。他当时认为宇宙应是静态的，于是人为地添加了一项“宇宙项”，以抵消动态变化获得静态解。但后来，爱因斯坦发现他自己犯了“一生中最大的错误”，因为美国天文学家哈勃发现宇宙是膨胀的而不是静止的，正如爱因斯坦添加“宇宙项”之前的引力场方程所预言的那样。有趣的是，后来爱因斯坦又犯了另一个错误，他发现自己构造的引力场方程还预测了一种直观上难以想象的情形——物质会在引力作用下发生坍塌，即引力场方程存在“奇点”解。他认为这是不可能的，但后来发现黑洞正是这种“奇点解”所预示的物理实在。这些经验使得当代物理学家们开始对数学上的结果持认真态度，以至于物理学家们如是写道：无论宇宙学中的数学作了何种预言，也不管这种预言对于我们有限的经验来说听起来是多么地不可能，我们都必须对之加以认真的考虑（戴维 •费尔津，赵复垣译，2000，P241）。事实上，不仅是宇宙学，在物理学中的其他领域也有许多类似例子，足以说明数学具有的不可思议的预测力，如狄拉克通过将相对论引入量子力学方程而发现正电子的预言，后来不久由威尔逊在迴旋加速器上果真发现了带正电的电子。事实上，科学思考的艺术——无论在物理学中、生物学中，还是在经济学中——就是决定作出什么样的假设。例如，假设我们从楼顶扔下来的是气球而不是大理石。我们的物理学家就会意识到，没有摩擦的假设在这种情况下并不正确：摩擦对气球的影响力要比对大理石大得多。重力在真空中发生作用的假设对研究大理石的下落是适用的，但对研究气球的下落并不适用。但是，当你问一个物理学家，一块大理石从 10 层楼的顶端掉下来需要多少时间，他会通过假设这块大理石在真空中落下来回答这个问题。当然，这个假设是不现实的。事实上，楼房周围是空气，空气对下落的大理石产生摩擦并使下落变慢。但物理学家将正确

地指出，这种对大理石的摩擦如此之小，以至其影响可以忽略不计。假设大理石在真空中下落使问题大大简单化，而又对答案没有实质性影响。因此，通过假设可以完成对模型的简化。针对不同的问题，可以采用不同假设。假设我们想研究当政府改变流通中的货币量时经济中会出现什么情况。这种研究的一个重要内容是价格会作出什么反应。经济中的许多价格并不经常变动。例如，报摊上的杂志价格若干年才变动一次。了解了这一事实后，当我们研究政策变动在不同时间长短中的影响时就作出不同的假设。为了研究这种政策的短期效应，我们可假设所有价格完全是固定的。但为了研究这种政策的长期效应，我们可假设所有价格完全是有伸缩性的。正如物理学家在研究大理石下落和气球下落时用了不同的假设一样，经济学家在研究货币量变动的短期与长期效应时也用了不同的假设。正是由于经济现象的高度复杂性，使得经济学家们在研究中不得不通过诸多假设来限定所考虑的研究范围和简化模型，只有这样才可能获得明确的结论和预测。但是，过多的假设也往往限制了理论的应用性，使得某些经济学理论显得华而不实，对现实缺乏直接的指导性和可用性。有这么一个笑话，就是讽刺经济学家在研究中过多使用假设来简化分析的弊病——据说有两位先生驾驶一架飞机翱翔于南美上空，但由于飞机失事，两人被困于安第斯山上。尽管不幸之中的万幸是俩人居然安然无恙，但却苦于被困在一个山顶上而找不到下山之路。忽然，他们瞧见山下有位牧羊人，于是便大声呼救，寄希望于牧羊人告诉他们下山的道路。牧羊人回答道：假设你们有两只降落伞，你们就可以从山顶上跳下山来。闻此牧羊人之言，俩人面面相觑，不约而同地说道：牧羊人想必是一位经济学家。

在不同的学术研究领域，通常会使用不同表现形态的模型。一般地，理论在本质上就是模型。但在工程学中，通常用实物材料做成模型，如建筑模型、生物力学模型、沙盘水利工程模型等，但在物理学和经济学中，常用的是数学模型。

模型的功用在于科学家们可以用它对即将发生的事件进行预测或者用它对已经发生的现象加以诠释。对于科学理论来说，其是否具有预测的功能被作为衡量理论是否“科学”的标准。理论是否“科学”与理论是否“正确”是不同的概念。“科学”的理论可能最终被证明是“不正确”的，而“正确”的理论却可能是“不科学”的。譬如，我们给出这样一个十分简单的“理论”；明天可能下雨也可能不下雨——这个“理论”完全是“正确”的，因为它在任何情况下（即“明天下雨和不下雨”）都是正确的，但我们很难说这样一种“理论”有什么具体用处，除了完全是一句废话外根本不具有任何价值。相反，“明天要下雨”这一“理论”完全可能被证明是错误的——只要明天不下雨——但它却是“科学”的，因为它含有某种由预测者携带的信息，即预测者所以作出这种预测必定是掌握了一定的信息。所以，只有含有信息量的理论才是科学的，否则没有任何用处。为什么“科学”的理论不等于是“正确的”理论呢？这是因为，科学理论一般是“全称”命题，即描述在普遍情况下都一般成立的规律（“规律”一词的含义就含有一

般性），即在所有相关条件下都成立的某种结论。譬如，“所有天鹅都是白的”这一“理论”就要求不仅现在存在的天鹅是白的，而且还要求过去曾经存在过的天鹅以及现在还未出现但未来可能出生的天鹅都是白的。所以，为了证明理论是否正确，就要求我们将所有的天鹅（包括现在还未出生的天鹅）都拿来看看是否是白的。显然，这是做不到的，因为未来才出现的天鹅现在还未出生呢？因此，我们在原则上不能做到能证明“所有天鹅都是白的”这一“理论”的正确性。早期的科学哲学主流学派即逻辑实证主义的哲学家们认为，理论是否科学要看理论是否能被实践证明是正确的，即理论是否能被验证。但是，实践总是有限的，而全称命题的实践证明要求完成无限次实践，这种矛盾导致有限的实践不能证明无限的理论的正确性。从逻辑学上看，实践是由单称命题组成的有限归纳过程，而理论是以全称命题表现出来的，逻辑学的基本常识告诉我们，单称命题的有限组成不能证明全称命题的正确性。有人会说，实践能证明理论的正确性是由人类既往的实践所证明了的“真理”，但是，这种说法犯了逻辑上的自我循环错误，因为它只不过是用待证明的命题去证明待证明命题，是用归纳法去证明归纳法的正确性，因为怎么能说既往的实践能证明实践能证明理论的正确性呢？因为既往的实践也是有限的呢（人类的历史是有限的，地球上生命的历史只不过 40 亿年，而整个宇宙的历史也才 150 亿年）。归纳法不能证明归纳法自身——这是著名的“休谟二难”问题（hume dilemma）。譬如，我们不能通过运用过去太阳每天都从东方升起的经验而作出明天太阳也一定会从东方升起的预言，因为从“过去”在逻辑上是不能推出“未来”的。有人会说，如果利用牛顿力学原理，人们能够从“过去”推出“未来”，但在逻辑上，我们怎么能保证牛顿力学在未来也必定成立呢？你可以说牛顿力学在过去被证明是正确的，但这只不过又是试图用归纳法去证明归纳法而已（事实上，牛顿力学在量子论和相对论出现后已被证明是“不正确”的）。一般地，无论是运用什么样的自然科学理论，我们在逻辑上不能从“过去”推出“未来”，量子力学和相对论或许有一天也会被新的理论所取代。休谟曾为我们讲了这样一个故事：一只小鸡在主人精心饲养下逐渐长大，小鸡发现每当主人向他伸出一只握有米粒的手时，它就可以饱餐一顿。因此，小鸡自然运用“归纳法”得出一个“理论”：只要主人伸手，她就可以吃到米粒。但是，有一天主人向它伸手，当它伸长脖子准备啄米时，主人一刀砍掉了她的脖子。小鸡不仅没有美餐一顿，反而成了主人的佳肴。显然，小鸡成为了“归纳法”的牺牲品。由于理论是全称命题，而实证却是有限的，实证不能证明全称命题的正确性，所以，我们在原则上不能证明理论的正确性，因为将理论的科学性定义为理论的正确性在研究中是不可操作的。事实上，1950 年代在澳洲就发现了黑天鹅——过去并不能证明未来！

那么，我们应该如何定义理论的科学性呢？显然，我们不能说因为某某理论符合另一权威理论的教条就认为其是科学的，因为即使是那一权威理论也至多经受了有限次归纳法的证实，其自身的科学性都还未解决，怎么能用它去证明别的理论是否科学呢？全

称命题的理论在逻辑上不能被单称命题的实证所证实，但同样从逻辑上看，单称命题的证伪却能够一举证明理论的缪误性，而单称命题的证伪只需一次性完成，这就为我们寻找衡量理论的科学性标准打开了一条思路——这正是剑桥哲学家波普尔的证伪主义哲学的逻辑出发点。当年波普尔正苦于难以寻找检验理论是否科学的准则时，他在普林斯顿大学拜访了爱因斯坦。当他得知爱因斯坦正在研究统一场论——试图将所有存在于大自然的四种力（引力、电磁力、强相互作用力和弱相互作用）——统一为一种基本力时，他问爱因斯坦："你是否认为你提出的这一理论就是正确的"。出乎他的意料之外，爱因斯坦回答说："我从来就不会认为我所提出的理论一定正确，我只不过等待着新的理论去取代它们"。听罢爱因斯坦教授的回答，波普尔顿时恍然大悟，随即提出了著名的"证伪主义"科学哲学体系——即可能被证伪的理论才是科学的，那些永远不能被证伪的理论是不科学的。因此，"明天要下雨"的"理论"可能被证伪（只要明天下了雨），因而是科学的，而"明天可能下雨也可能不下雨"的"理论"永远不会被证伪，所以不是科学的。波普尔的证伪主义哲学不仅对后来的科学哲学发展起到了巨大的推动作用，而且还对科学家的思维产生了重要影响。

许多当代科学家都一致认为证伪主义哲学所描述的正是科学家通常所进行的思维方式。对现代经济学家来说，波普尔的思想更是居主流地位。弗里德曼曾说：波普尔哲学所说的正是我长期所采用的思考方法。

为什么理论是否科学取决于其是否可能被证伪呢？这是因为，只有可能被证伪的理论才具有预测功能，才包含有可用的信息。"明天可能下雨也可能不下雨"的"理论"等于什么也没说，不具有任何的预测价值，完全是废话，因而不是科学的"理论"。由于我们在原则上不能"证实"理论，因而波普尔之前的逻辑实证主义提出的"理论被证实就是科学的理论"的主张是没有价值的，所以，我们不能沿着理论是否可能被"证实"的思路去建构判定理论的科学性的准则。

我们所能够做的就是退后一步——不去试图证实理论的正确性——而是去看是否有更好的理论取代现有理论。当我们还没有找到更好的理论取代现有理论时（更好的理论是具有更好预测能力的理论，如果一个新的理论能预测原来的旧理论不能预测的事件，那么，新理论就是比旧理论"更好"的理论），我们就没有理由放弃现有的理论。同样，我们使用现有的理论并不意味着我们完全相信这个理论就是正确的，只不过我们还未找到一个更好的理论取代它罢了！我们相信现有理论在未来会被新的理论所取代，但在此事发生之前，我们不得不继续使用现有理论。正如爱因斯坦所说，我们正等待着现有理论被推翻，从而完成从旧理论向新理论的过渡呢？这正是科学革命或理论进化的逻辑。根据这种科学哲学的准则，经济学（它正试图成为基于逻辑与实证基础上的现代科学的一员）理论是否科学也应看其是否具有预测功能。经济学家斯蒂格勒曾批评张伯仑提出的"垄断竞争"理论不具科学价值，因为"垄断竞争"理论所能作出的预测由已

有的完全竞争和完全垄断理论都能作出，而在某些由完全竞争和完全垄断理论可以作出预测的场合，由垄断竞争理论却不能作出预测。譬如，当税率上升后，完全竞争和完全垄断理论可以对产量和价格变动作出预测，但垄断竞争理论却不能作出相应的预测。因而，从证伪主义角度看，从完全垄断和完全竞争理论过渡到垄断竞争理论不能被视为是一种经济学理论发展上的进步，垄断竞争理论也不能被视作是“科学的”理论，尽管垄断竞争理论的提出完全是基于其假设比完全竞争和完全垄断理论的假设更为现实。这样，我们看到了模型假设过于贴近现实反而使模型降低了预测能力和其“科学性”的一个具体例子。不幸的是，目前博弈论似乎也存在类似的问题，因为许多博弈论模型存在多重纳什均衡，使得其预测功能模糊。但在博弈论场合，这个问题还未有定论，博弈论专家们正在不断努力，通过引入更多的均衡概念去消除多重均衡，尽管这方面努力已取得了不少成果，但离最终目标——只留下唯一的一个均衡——还有待时日。但是，目前还没有证据表明这一目标的实现存在着什么不可逾越的障碍。事实上，我们将在第四章中说明，博弈论专家最近在此方面已经向前迈进了一大步，甚至有可能离此最终目标仅有咫尺之遥了！

2.1.3 经济学与经济模型

现代经济学中的模型化方法主要是基于数学模型的推演和分析，而经济学研究运用数学模型的鼻祖大概要追溯到 17 世纪英国经济学家威廉·配弟那里，但数理经济学（Mathematical Economics）的真正先驱是 19 世纪的法国经济学家古诺（Cournot）。有趣的是，正是在古诺的市场竞争模型中可以找到博弈论思维方式的影子。事实上，所谓“古诺博弈模型”已成为博弈论教科书中常见的例子。经济学大规模采用高等数学方法始于 1870 年左右，当时奥地利学派的门格尔、英国剑桥大学的杰文斯和瑞士洛桑大学的瓦尔拉倡导用微积分方法研究经济行为，随即便出现了所谓“边际学派革命”，新古典经济学由此萌芽。边际学派以瓦尔拉、门格尔和杰文斯为首领，开始通过运用数学模型研究诸如价格和产量这些经济变量的决定问题。瓦尔拉是边际学派的第一领头人，他的贡献在于提出了一般均衡的思想。瓦尔拉在其名著《纯粹国民经济讲义》中，开始尝试用数学方程组来平衡各种商品的供求，并指出方程组的解就决定了实际会出现的商品价格和数量。尽管瓦尔拉当时并没有证明方程组存在解以及解的唯一性，但他实际上提出了对后世经济学发展影响极大的一般均衡理论。我们没有理由责怪瓦尔拉在数学上的不严密性，因为在 19 世纪还未出现可以用来证明一般均衡解的存在唯一性的数学方法，而这要等到 20 世纪 50 年代才由数理经济学家德布鲁（Debreu，1959）完成这一工作（德布鲁因此于 1983 年与肯尼思·阿罗一起获得诺贝尔经济学奖）。

到了 20 世纪初期，新古典经济学才由剑桥学派创始人马歇尔集其大成。马歇尔本是粒子物理学博士，当他取得博士学位后，他没有继续研究当时正处于科学革命热潮中

的微观物理学，而在物理学发展历史上激动人心的年代里改行进入经济学领域。据说马歇尔博士作出这一选择是基于其打算寻求济世良方的愿望，而经济学正是他所寻求的目标。马歇尔将物理学的思考方法引入经济学，并基于科学主义的精神将当时常用的“政治经济学”名称改为“经济学”，还写出了闻名于世的教科书《经济学》。

马歇尔发现瓦尔拉的一般均衡模型在数学上十分复杂，于是将其简化为“局部均衡”模型。在一般均衡模型中，不同商品和要素市场的价格和数量之间存在着互动关联（由于存在互补品和替代品），因而所有商品市场的价格和数量都会同时发生变动，所以分析难以得出明确的结论，影响了模型的预测功能——这正是模型的复杂性影响了模型的预测能力的一个具体例子。为了获得预测功能，马歇尔对瓦尔拉的一般均衡模型进行简化。他将单独一种商品市场从其它商品市场中分离开来，即给定其他商品市场处于均衡状态的假定，单独研究某一商品市场的价格和数量的决定。基于物理学训练的背景，马歇尔将物理学中的“平衡”思想引入经济学。在物理学中，一个静止的物体并不一定意味着它没有受到力的作用，而更多的是意味着物体受到来自相反的各种力量的作用且这些作用相互之间恰好抵消。譬如，静置在桌面上的茶杯受到向下的重力的作用，也受到桌面向上的反作用力的作用，且它们之间恰好相互抵消才使得茶杯在桌面上处于静止状态。即使在宇宙深处远离其他星球的一个静止物体，它也受到来自各个方向的遥远星球的引力作用，只不过这些力相互抵消罢了。因此，处于静止状态的物体并不能逃避力的作用，只不过这些作用互相抵消掉了。这种思想的重要性在于——当某些力的作用减弱或消失或增强时，其他力的作用效应就会马上显现出来，物体就不会继续处于静止状态。当我们对物体的受力情况进行分析之后，我们就能预测某些力的作用发生变化后物体的运动趋势。马歇尔深受物理学中这种思考方法的影响，他成功地将其移植于经济学研究中。他假定，经济即使处于静上状态，也并不意味着不存在经济力量的影响，只不过这些影响相互抵消掉了而已。当某些经济力量的作用减弱或消失或增强时，经济就会进入变动时期。只要弄清楚了这些经济力量的作用机理，就能预测经济力量的相对强弱发生变化后经济状态的演化方向。马歇尔认为最重要的经济力量是供给与需求。他在局部均衡模型中对供给与需求进行了深入研究，并考虑当这两种力量相互平衡时经济的静止状态的主要特征。他将物理学中的上述“平衡”概念在经济学中的类比概念称为“均衡”（equlibrium）。按马歇尔的意思，“均衡”就是指影响经济状态的各种经济力量相互抵消的情形（在瓦尔拉那里，尽管他最早提出了一般均衡的思想，但还没有使用“均衡”的术语）。马歇尔认为最重要的经济力量是供给与需求，他用供求两种力量的对比和平衡研究了市场均衡问题，并给出了均衡价格和均衡数量的决定方式。局部均衡指给定其他商品市场供求已达到均衡的情况下，某一商品市场的均衡状态。当改变其他商品市场的均衡位置时，某种商品市场的均衡位置也会发生相应的变化，这就是局部均衡比较静态分析。马歇尔强调供求力量的互动对经济均衡的作用，对后来的经济学发展影响甚大，

以至于萨谬尔逊在其著名的《经济学》教科书中写道：“只要鹦鹉能说供给与需求，她就成为了经济学家”。从科学方法论上看，马歇尔的最大贡献莫过于假设“均衡”就是实际会出现的状态，尽管这种假设的现实基础不甚牢靠，但它却明确标明了模型的预测功能——即将均衡作为对实际经济现象的预测。

尽管马歇尔的局部均衡模型可以对经济行为作出预测，但这种方法仅在存在许多相互竞争的经济行为主体情况下才是充分有效的，即其预测结果才是较为准确的。相反，当经济行为主体很少时，譬如寡头垄断市场上的厂商数目就很少，则由局部均衡模型给出的预测与实际情况之间的偏差就很大，这是由于局部均衡中的假设忽略了经济行为主体之间的互动。在经济行为主体很少时，经济行为主体在进行决策时要考虑到其他经济行为主体对其各种行为选择的预期反应，从而在效用最大化决策中需要纳入这种反应，因而行为主体之间的决策存在相互依赖性。行为主体的效用函数不仅是其自身选择行为的函数，而且也是其他行为主体行为选择的函数。譬如，对于寡头垄断市场上的寡头企业，它们在决定其价格或产量选择时需要同时考虑这种选择对其他寡头企业的影响，考虑其他寡头企业在它作出决策之后会如何改变它们的价格和产量，而这个寡头企业的利润不仅是其自身价格和产量决定的函数，而且还是其他寡头企业价格和产量决定的函数。在著名的古诺型寡头产量博弈模型中，给定另一个寡头厂商的产量，一个寡头厂商可选择其利润最大化产量，但给定该寡头厂商生产其选定的利润最大化产量，另一寡头厂商也会选择其自身的利润最大化产量，而这一产量可能不等于开始给定的产量。这一产量调整过程可能会不断进行下去，直到达到均衡为止。在均衡状态，任一寡头厂商的产量都是给定另一寡头厂商产量条件下的利润最大化产量，从而不再出现互动。在马歇尔的局部均衡模型中，如果假定厂商有无限多个，消费者也有无限多个，就可以不考虑任一厂商或消费者在作出决策时对别的厂商或消费者决策的影响，这样就略去了经济行为主体之间的复杂互动过程，使分析变得较为简单，这就是马歇尔的完全竞争模型。在另一种极端情形，倘若假设只有一个厂商，则得到马歇尔的完全垄断模型。张伯仑和罗宾逊夫人为了在这两种极端情形之间插入一种过渡型模型，于 20 世纪 30 年代提出了垄断竞争模型，但该模型仍然假定经济行为主体有无限多个，从而略去了经济行为主体间的互动。尽管在早期的微观经济学中，寡头垄断市场仍然是一个重要内容，但是这部分内容总是以一种十分特别的形式出现的，即区别于其他都存在局部均衡位置的市场结构如完全竞争市场、完全垄断市场和垄断竞争市场，寡头垄断市场却没有其一般形式的局部均衡位置。在早期的微观经济学教科书中，寡头垄断市场行为表现为一系列基于对寡头行为机制不同假定的模型，而这些模型对寡头行为的预测是不同的，这种不同完全是由于模型中对寡头间的互动机制的不同假定。从经济思想角度看，正是这些不同的模型

蕴酿出了后来的博弈论思想。在经济学家开始用博弈论系统地刻划寡头行为之前，描述寡头行为的这种考虑了互动过程的模型在微观经济学中只是一种补充性的内容，当时在研究上主要还是采用马歇尔的略去了经济行为主体间互动过程的局部均衡模型。但是，随着这类纳入了互动过程的模型的不断出现，经济学家开始意识到互动的重要性，并且发现马歇尔的方法其实只是这类模型的一种特例，即当纳入互动的模型中经济行为主体数目趋于无限大时，这类模型就自然给出了经马歇尔方法处理的结果。因为随着经济行为主体数目的增加，经济行为主体间的互动就显得不太重要了，当相互影响的经济行为主体数目达到无限大时，这种互动效应就完全消失了。随着博弈论的出现，经济学家们逐渐有了系统处理经济行为主体间互动过程的工具，他们发现可以将经济分析拓宽到更有效处理经济行为主体数目较少的情形，而原有的马歇尔式简化方法可以作为一种特例。这样，经济学在方法论上获得革命性进步，因为博弈论将经济分析的效率提高了，将经济分析的有效范围也拓宽了。在 Tirole 的名著《产业组织理论》（马捷等译，1997）中，作者使用博弈论方法重新完成了产业组织理论这一传统上主要是基于文字描述性的学科在方法论上的更新，使产业组织理论从一个较为冷僻的学科一跃成为一个十分活跃的经济学前沿研究领域。Kreps 的教科书《微观经济学教程》（Kreps，1990）在寡头垄断内容部分全部用博弈论处理，使得该书迅速蜚声国际经济学界。在 1940 年代末和整个 1950 年代，塔克教授在普林斯顿大学讲授博弈论，举办博弈论讨论班，从而培养出一大批掌握了博弈论方法的经济学家，这些深谙博弈论方法的学者群被戏称为“塔克军团”，普林斯顿大学也因此博得了“博弈论工厂”的美名。或许正是由于这些普林斯顿工厂的大批“产品”在国际经济学界开始产生影响，博弈论于 1970 年代以后风靡经济学各个领域，并逐渐上升到作为主流经济学的基本方法论的地位。

在博弈论发展历史上，纳什（Nash）的创造性工作具有划时代的学术意义。在博弈论创始人冯 • 诺依曼和摩根斯坦提出最早的博弈论构架时，他们还不能很好地处理局中人之间的互动。在这种情况下，他们提出的均衡概念是没有充分纳入局中人间互动机制的“最大最小”均衡，即局中人选择使其最小支付或效用最大化的策略。当所有局中人都如此完成各自的决策时，倘若所有局中人都不再调整其选择，则博弈就达到“最大最小”均衡。但是，冯 • 诺依曼和摩根斯坦的“最大最小”均衡没有考虑到局中人之间在策略选择上的互动，其预测能力是要大打折扣的。譬如，我们考虑表 2.1 中的双变量矩阵（也称“支付矩阵”）所描述的二人博弈。我们用 A 和 B 分别代表两个局中人，用 U 和 D 表示 A 的二个可选择行动，L 和 R 表示 B 的两个可选择行动，矩阵中的每一小方格中的一双数字分别表示当 A 和 B 正好选择该小方格对应的行动时各自所获得的支付，习惯上小方格中左边第一个数字表示 A 的支付，第二个数字表示 B 的支付。

表 2.1　二人博弈的支付矩阵

		B: L	R
A	U	3, 2	2, 3
	D	4, 5	3, 3

我们先通过“最大最小”方法找出博弈的均衡解。当 A 选行动 U 时，A 可能获得的最小支付为 2（当 B 选 R 时）；当 A 选行动 D 时，A 可能获得的最小支付为 3（当 B 选 R 时）。这种方法认为 A 不知道 B 会选 L 还是选 R，所以为了保险，假定 A 会选使其在最不利的情况下的支付最大化的行动，所以，该方法预测 A 会选 D。根据同样的逻辑，该方法预测 B 可能会选 R，于是“最大最小”均衡为：A 选 D，B 选 R，用（D，R）表示，括号中第一个符号表示 A 的均衡行动，第二个符号表示 B 的均衡行动。

但是，局中人 A 和 B 果真会按上述预测的均衡进行选择吗？让我们从另一个角度来看看另一种预测方法。我们纳入 A 和 B 之间的互动。显然，对 A 来说，无论 B 是选 L 还是选 R，A 选 D 的支付都大于选 U 的支付，所以，A 不会选 U，我们称行动 D 是占优（dominant）于 U 的[1]，或称 D 是占优战略（我们称局中人的一个选择为一个“战略”）。显然，A 只会选择占优战略 D。然后，我们再来看 B。根据理性人假定，B 也知道 A 会选 D，因为假若他是 A，他就会如 A 那样选择行动 D。接下来，给定 A 选 D，B 当然会选 L。于是，均衡不会是（D，R）而是（D，L）。这种预测方法被称为“剔除劣战略”的方法，所得到的均衡被称为“剔除劣战略的占优战略均衡”。

那么，到底哪一种方法得到的均衡更为合理呢？哪一种方法的预测更为可信呢？答案依赖于我们对局中人的理性假定。倘若我们假定局中人是充分理性的，他们就会按“剔除劣战略”的方法去进行博弈，因而“剔除劣战略的占优战略均衡”是一种更为合理的均衡，而由这种均衡给出的预测就更为令人可信。事实上，由“最大最小”方法得到的“均衡”很难说是一种“均衡”。给定 A 选 D，B 应选 L 而非 R。

倘若所有局中人都可同时剔除各自的劣战略而达到均衡，则称该均衡为“占优战略均衡”。在博弈论中，含有占优战略均衡的一个著名例子是由塔克给出的“囚徒困境”（prisoners’ dilemma）博弈模型。该模型用一种特别的方式为我们讲述了一个警察与小偷的故事。假设有两个小偷 A 和 B 联合犯事、私入民宅被警察抓住。警方将两人分别置于不同的两个房间内进行审讯，对每一个犯罪嫌疑人，警方给出的政策是：如果一个犯罪嫌疑人坦白了罪行，交出了赃物，于是证据确凿，两人都被判有罪。如果另一个犯罪嫌疑人也作了坦白，则两人各被判刑 8 年；如果另一个犯罪嫌人没有坦白而是抵赖，则以妨碍公务罪（因已有证据表明其有罪）再加刑 2 年，而坦白者有功被减刑 8 年，立即释放。如果两人都抵赖，则警方因证据不足不能判两人的偷窃罪，但可以私入民宅的

罪名将两人各判入狱 1 年。表 2.2 给出了这个博弈的支付矩阵。

表 2.2 囚徒困境博弈

		B 坦白	B 抵赖
A	坦白	–8, –8	0, –10
A	抵赖	–10, 0	–1, –1

我们来看看这个博弈可预测的均衡是什么。对 A 来说，尽管他不知道 B 作何选择，但他知道无论 B 选择什么，他选择“坦白”总是最优的。显然，根据对称性，B 也会选择“坦白”，结果是两人都被判刑 8 年。但是，倘若他们都选择“抵赖”，每人只被判刑 1 年。在表 2.2 中的四种行动选择组合中，（抵赖、抵赖）是帕累托最优的，因为偏离这个行动选择组合的任何其他行动选择组合都至少会使一个人的境况变差。不难看出，“坦白”是任一犯罪嫌疑人的占优战略，而（坦白，坦白）是一个占优战略均衡。

但是，现在的问题是我们能否保证博弈模型都存在“占优战略均衡”呢？不幸的是，答案是否定的，许多博弈都不存在“占优战略均衡”。下面再看表 2.3。

表 2.3 不存在占优战略均衡的二人博弈

		B L	B R
A	U	4, 3	2, 5
A	D	3, 8	6, 7

显然，A 和 B 都没有占优战略。当 B 选 L 时，A 选 U 获得的支付比选 D 获得的支付大，但当 B 选 R 时，A 选 D 获得的支付比选 U 获得的支付要大。类似地，当 A 选 U 时，B 选 R 获得的支付比选 L 获得的支付要大，当 A 选 D 时，B 选 L 获得的支付又比选 R 获得的支付要大。因此，A 和 B 都没有占优战略。

那么，在一般情况下，博弈模型是否存在可用于预测的“均衡”呢？这显然是博弈论首先要予以回答的基本性问题。在这里，“均衡”概念自身也不是预先给定的，而是需要根据合理的（基于理性人假定的）逻辑步骤加以导出的概念。为博弈论作出基础性贡献的纳什在 1950 年代完成了一系列创造性研究工作，他提出的“纳什均衡”在非常一般的情况下都存在。这种一般结论是指当局中人数目和每个局中人可选择的战略数目都是有限的时候，博弈模型至少存在一个纳什均衡（此时，均衡还包括我们将在下一章中将要介绍的混合战略博弈均衡）。我们称局中人可选择的战略的范围为“战略空间”。

当战略空间中的可选择战略数目为有限数时，称战略空间是有限的。我们称局中人数目和每个局中人战略空间都是有限的博弈为“有限博弈”。纳什的研究保证每一个有限博弈都至少存在一个纳什均衡。这个结果被称为“纳什存在性定理”。在纳什之后，德布鲁（Debreu），Glicksberg, Fan 等人又给出了纳什存在性定理的几种不同形式和在一定程度上对其加以了扩充。不幸的是，纳什存在性定理不排除存在多重均衡的情形，因而博弈模型的预测能力仍处于不尽完善之中，尽管泽尔腾（selten）、kreps 等人在动态博弈方面给出了更为精致的均衡概念，但多重均衡问题至今仍未解决。尽管如此，博弈论自纳什开始取得了长足进步。纳什的贡献不仅仅是证明了纳什存在性定理，更重要的是他给出了我们现在进行博弈论研究所使用的一整套方法和工具箱，他为非合作博弈论提供了一个完整的思维框架和学术体系（当我们假定局中人都是不受某种强制性协议约束的效用最大化追求者，博弈就是非合作博弈。现在经济学家谈到博弈论，一般就是指非合作博弈，很少指合作博弈。合作博弈与非合作博弈之间的区别主要在于人们的行为相互作用时，局中人能否达成一个具有约束力的协议。其实，从根本上看，只存在一种博弈即非合作博弈，就象萨特所说的：人是自由的，局中人本质上总是可在许可范围内自由选择其战略。只不过在合作博弈场合，违反协议带来的局中人个人成本可能是无限大而已）。

表 2.4　存在一个纳什均衡的二人博弈

		B: L	B: M	B: R
A	U	3, 1	4, 5	1, 6
	D	7, 8	2, 7	2, 3

所谓纳什均衡，是指局中人在选定某一战略组合时，每一个局中人都没有积极性改变他已选定的战略，除非其他人会改变他们已选定的战略，但由于其他局中人也同样没有积极性改变已选定的战略，所以，局中人都没有积极性偏离已选定的战略，这种“僵局”显然可作为一种均衡概念加以定义。在纳什均衡中，局中人的“最优”战略取决于其他局中人的战略选择。也就说，局中人选择的战略是否“最优”取决于其他局中人选择什么样的战略。当每个局中人选择的战略在给定其他局中人的战略选择时都是最优战略时，这时的战略组合就构成一个纳什均衡。

在表 2.4 给出的二人博弈中，局中人 A 和 B 都没有占优战略，从而没有占优战略均衡，但它却有一个纳什均衡。给定 A 选择 D，B 选择 L 是最优的选择；反过来，给定 B 选择 L，A 原来的选择 D 是最优的选择。因此，纳什均衡为 A 和 B 的战略组合（D，L）。

2.1.4 经济现象与博弈模型

经济学家使用博弈论的目的是试图通过博弈论沟通许多表面上看来完全不同和相互之间似乎完全没有联系的诸多经济现象（甚至包括政治的、军事的、社会的和生物学上的各种现象），从而完成一种可解释不同经济现象的一般性理论或机制体系的建构，通过博弈论模型抓住诸多现象的本质机理。这种想法在实用上是为了完成一种对事物产生机制进行分类，从而能使人们更为简化和系统地分析、理解各种经济现象甚至社会、军事、政治和生物学上的各种现象的发生机理和对现象结果加以预测的方法，当博弈论经济学家面临需要加以解析的经济现象时，他首先需要做的事就是看看该现象的发生机制是否可归于某种已知的博弈论模型。倘若可归入一个已知的博弈论模型，他就立即知道现象的本质和发生机制，也就立即可对经济现象的结果进行预测。如果某位经济学家发现一种新的现象或机制不能归入已知的所有博弈论模型，则他会因此而感到激动和兴奋，因为他会因此而建构新的博弈模型去解释这一全新的现象。如果他能成功地做到这一点，那么，他就在博弈论经济学中作出了创造性贡献。当然，即使能将现象归入已知的某一类博弈模型，往往也在开始不是一眼就能看出来的，而是要通过复杂的分析研究才能完成这种归类，这种工作也是极有价值的，也仍然具有重要的学术意义。

譬如，我们在前面所介绍的“囚徒困境”博弈模型实际上就刻划了许多表面上看来是完全不同的诸多现象的本质。如公共品的提供、卡特尔的不稳定性、军备竞赛、团队生产的低效率、改革激励的不足等等。当商品市场上的两个寡头企业进行产量博弈时，如果两个企业联合起来形成卡特尔，选择垄断利润最大化的产量，每个企业都可以得到更多的利润。但卡特尔协定不是一个均衡，因为给定对方遵守协议的情况下，每个企业都想增加生产，结果是每个企业都只得到“囚徒困境”均衡中的利润支付，它严格小于卡特尔产量下的利润。这个例子说明，有时个人理性与集体理论的冲突对社会来说也许是一件好事，尽管它对该集体的成员来说是一件坏事（显然，对于囚徒困境故事本身来说也是如此）。公共品的供给也是一个囚徒困境问题。如果大家都出钱兴办公用事业，所有人的福利都会增加。但事实上可能出现我出钱而你不出钱时我将得不偿失，而你出钱我不出钱可以占便宜的情况，此时每个人的最优选择都是“不出钱”，好事就做不成。在冷战期间，苏美两国都竞争增加各自的军费预算，如果不搞军备竞赛，各自把资源用于民品生产，不是很好吗？但问题是，如果我将资源用于民品生产，你增加军费时我就面临危险了。于是，均衡是两国都大量不断地增加军费开支和增加导弹生产量，即使多次签订限制战略性武器条约也不能竭制军备竞赛的势头，最终导致前苏联被拖垮。事实上，美国的战略家们正是基于这种囚徒困境逻辑，利用军备竞赛将前苏联经济搞垮。在经济改革中也存在囚徒困境现象。改革者要付出成本，而改革的成果由大家共享，于是，尽管大家都认为改革是好事，却没有人有积极性去全力投入改革。根据博弈论的观点，一种制度安排若要发生效力，必须是一种（纳什）均衡。否则，这种制度安排便不能实

施。在 1970 年代初的第三次中东战争结束时，由于美国等西方国家在战争中帮助了以色列，代表主要石油输出国集团利益的国际石油卡特尔 OPEC 通过提高油价对西方进行惩罚。当时，这一行动导致美国油价暴涨。对于象美国这样一个建立在汽车轮子上的国家来说，油价上涨三倍，对其经济和社会的影响是十分巨大的。许多人排长队购买配给的汽油，一些工厂被迫减产甚至关闭，人们一片怨声载道声。当时，麻省理工学院的一位经济学家在报上著文指出，油价暴涨不会持久，待一段时间过去后油价会自然回落，而这段时间不会很长久。果不其然，当高企的油价持续不太长的时间后，就下跌到较低水平。这位经济学家有什么特别的本事能预测到这种结果呢？其实内中并无什么特别奥妙，他只不过是简单地运用了一下“囚徒困境”模型而已。OPEC 是一个国际卡特尔组织，主要由包括沙特阿拉伯、伊朗、伊拉克、科威特、突尼斯、尼日利亚、突尼斯、委内瑞拉等 12 个主要石油生产和出口国组成。“囚徒困境”模型告诉我们，合作行为不是均衡，而均衡是不合作的。OPEC 中的不同国家间经济社会条件存在着很大差异，一些国家（如沙特阿拉伯、科威特等）经济实力较强，能承受因对西方的制裁而减少石油出口带来的损失，但其他一些经济实力较弱的国家（如突尼斯、委内瑞拉等）就不能承受因制裁西方而减少石油出口的损失。这样，经济实力较弱的国家就有私下不遵守卡特尔共谋协议的动机而缺乏遵守协议的积极性。这些国家的政府就可能对民间石油出口的违反协议行为（如生产和出口的石油桶数超过了 OPEC 的规定）睁只眼闭只眼，而给定这种可能性，其他包括那些经济实力较强的成员国也就产生了违反卡特尔协议的动机。因为预期到某些成员国会违反协议，协议也就丧失了预期效力，而给定协议丧失效力，自我约束就不是最优选择；相反，此时的最优选择是赶紧抢在别人多销石油的前面多销自己的石油。这样，无论是富国还是穷国，OPEC 中各个成员都有私下违反协议而增加石油出口的动机。于是，卡特尔协议规定的共谋在不长的时间内就崩溃了，各个成员国未能按协议规定的上限出口石油，而是更多地生产并出口石油，造成原来制造的高企油价很快下跌。麻省理工学院的那位经济学家就正是运用了我们在前面所说的方法，即首先将现象归入“囚徒困境”这一模型，然后利用该模型预测卡特尔的不稳定性，对最终结果提出预见性的判断。

另一个例子是中国的几个生产彩电的寡头厂商于2000年合谋将彩电价格维持高位，他们合谋搞了个所谓的“彩电厂家自律联盟”，并在深圳举行了由多家彩电厂商首脑参加的“彩电厂商自律联盟高峰会议”。当时，国家有关部门还未出台相关的反垄断法律，对于这种在发达国家明显属违法行为的所谓“自律联盟”，国家在法律上暂时还是无能为力的。寡头厂商在光天化日之下进行价格合谋，并且还通过媒体大肆炒作，而政府却对此无所作为，岂非咄咄怪事，这在发达国家是不可思议的。但是，尽管政府没有制止这种事情，公众也不必担心彩电价格会上涨。因为，正如当时中国的一些经济学家所指出的那样，“彩电厂商自律联盟”只不过是一种卡特尔行为，其内在的不稳定性和非合

作均衡特点将导致其自然破产，彩电价格不会上涨。果真如此，在“高峰”会议之后不到二周，国内彩电价格不是上涨而是一路下跌。因为卡特尔内所有成员都具有一种机会主义倾向：给定其他成员遵守协议进行提价或不降价，我自己偷偷降价不是可以更快地将别人的市场抢过来吗？而给定其他成员也作如是打算，他们也可能降价，那么，自己则更需要尽快以更大幅度降价才不会丢失已有市场。给定每一成员都会有这种想法，其结果是大家都会争相降价，最终导致价格不是上涨而是暴跌。事实上，当时在“高峰”会议之后，寡头们就对代理销售商的降价行为视而不见，而代理商又通过媒体明确表示不会提价，最后终于酿成一股降价之风。[2]

我们已经看到博弈论模型在预测上所具有的神奇威力。其实，这在博弈论经济学家那里只属平常本事，类似例子在博弈论经济学中可谓比比皆是。博弈论模型不仅可以用来预测现象，而且还可以用来对已观察到的现象进行解释。譬如，前面已经提到的“二手车”模型不仅可以用来解释二手车市场上的逆选择现象，而且还可说明一个单位的人才流失、政坛上的腐败蔓延等等。

注释：

[1] 田国强博士首先将 dominant 译为“占优”，某些教科书将该词又译为“优超”或“上策”等。

[2] 当然，在“囚徒困境”模型中，如果我们将支付数字改变，则均衡就可能也相应改变，或许不合作行为不再是均衡。但此时模型本身也变化了，新的模型不再可用来解释不合作现象但也许可用来解释其他别的什么现象。“囚徒困境”模型不是指这个故事本身，而是着重指不合作现象。

第三章　完全信息静态博弈

正如我们在第二章中所指出的那样，局中人在博弈中所享有的信息量对于博弈的结果有着重大的影响，博弈中的信息对称程度（或不对称程度）也决定着博弈的特征和结果。除此之外，博弈中各局中人在选择其行动时的先后顺序也决定着博弈的特征和结果。

譬如，有两个销售同样产品的销售商 A 和 B 打算进入某一区域性市场。由于这个区域市场对产品的需求是有限的，当他们都同时进入该区域市场时，他们各自占有的市场规模都偏小，从而造成 1 个单位的亏损；但是，当只有一个销售商进入该区域性市场时，则获得 1 个单位的利润；当然，不进入市场时的利润为零。假如 A 和 B 同时进行决策或者他们在进行各自的决策时并不知道另一方的选择，则博弈就被称为是一种“静态”博弈，刻划它们的支付情况的矩阵被称为“支付矩阵”，见表 3.1。

表 3.1　市场进入的静态博弈

		B：进入	B：不进入
A	进入	–1, –1	1, 0
A	不进入	0, 1	0, 0

A 和 B 的行动选择范围都是“进入”或“不进入”。当 B 选择“进入”时，A 的最优行动选择是“不进入”，而给定 A 选择“不进入”时，B 的最优选择是“进入”。因此，{不进入，进入}是一个纳什均衡。类似地，{进入，不进入}也是另一个纳什均衡。

下面，我们将这个博弈作一种修改，假定博弈是“动态”的，即 A 和 B 在行动选择上有“先”与“后”的顺序。假定 A 先选择，B 在 A 完成了其选择后再进行自己的行动选择，并且 B 在进行行动选择前知道 A 的选择结果。此时，我们用图 3.1 来表示这个博弈。

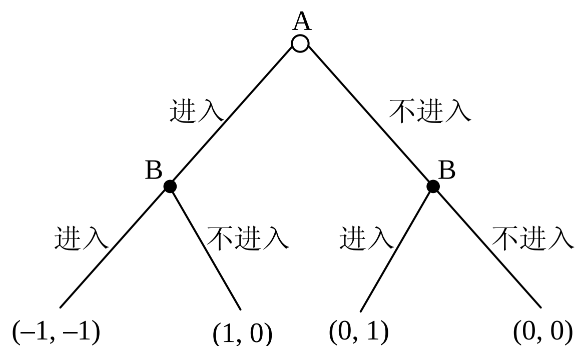


图 3.1　市场进入的动态博弈：A 先行动

在图 3.1 中，空心和实心的小圆点被称为决策结（decision nodes），位于决策结旁边

的字母代表在这个决策结处进行行动选择的局中人，该局中人在此决策结处进行行动选择。通常，整个博弈中进行第一个行动选择的决策结用空心圆点表示。图中的线段被称为“枝”（branches），一个枝表示位于该枝上端决策结处的局中人在该决策结可能选择的一个“行动”。最下方的枝的下端被称为终点结（terminal nodes），当博弈进行到任一终点结时，博弈过程就告结束。终点结处的向量表示博弈进行到此处从而结束博弈时局中人的支付，向量中从左端数起第一个数字是最先行动的局中人的支付，第二个数是第二行动的局中人的支付……等等。每一个枝旁边的文字指出了该枝代表的行动。博弈的行动顺序在图中是从“上”到“下”，现在，我们来看看这个博弈的“均衡”是什么。当 A 选“进入”时，B 接着就会选“不进入”；当 A 选“不进入”时，B 接着选的是“进入”。A 比较两种选择所获得支付，从而选择“进入”。于是，博弈只有一个“均衡”，即{进入，不进入}。均衡结果是：A 选“进入”，B 选“不进入”。图 3.1 中的几何图形常被称为“博弈树”。

显然，行动有先后之分的“动态”博弈与“静态”博弈相比会给出不同的预测结果。下面，我们再将上述“动态”博弈中局中人 A 和 B 的行动顺序作一下修改，即假设 B 先行动，A 在观察到了 B 的行动选择后，再进行自己的行动选择。我们用图 3.2 来表示这个博弈。

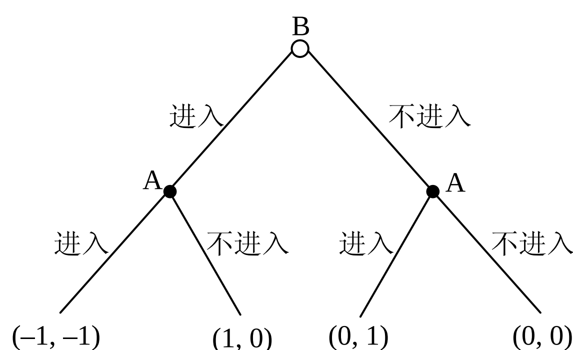


图 3.2　市场进入的动态博弈：B 先行动

显而易见，图 3.2 与图 3.1 所表达的博弈相比除行动的先后顺序不同外，其他结构特征完全相同。所以，均衡结果是：A 选“不进入”，B 选“进入”。这样，我们看到，博弈的行动顺序会影响博弈的预测结果。我们通常将支付矩阵（或代数形式）表达博弈的方式称为博弈的“战略式”表述（strategic form representation）或标准式表述（normal form representation），而将用图 3.1 和图 3.2 中那种“博弈树”表达博弈的方式称为博弈的“扩展式”表述（extensive form representation）。两种表述方法本质上是相同的，但通常用战略式表述表达静态博弈较为方便，而动态博弈用“扩展式表述”更为直观一些。譬如，我们用表 3.2 中的战略式表述来表示由图 3.1 给出的动态博弈。

表 3.2　用战略式表述表示图 3.1 中的动态博弈

		B			
		{进入，进入}	{进入，不进入}	{不进入，进入}	{不进入，不进入}
A	进入	–1, –1	–1, –1	1, 0	1, 0
	不进入	0, 1	0, 0	0, 1	0, 0

此时，需要注意的是，这时 B 的每一个选择或战略都必须给出任何情况下（无论其是否真正可能出现）自己的行动选择。在表 3.2 中，B 有 4 种选择或战略，每一个战略用一个向量表示，向量中左端第一个行动选择对应于 A 选“进入”时 B 的行动选择，第二个行动选择对应于 A 选“不进入”时 B 的行动选择。给定 A 选“进入”，B 的最优选择是{不进入，进入}和{不进入，不进入}。当 B 选{不进入，进入}时，A 的最优选择是“进入”，当 B 选{不进入，不进入}时，A 的最优选择是“进入”；当 A 选“不进入”时，B 的最优选择是{进入，进入}和{不进入，进入}；当 B 选{进入，进入}时，A 的最优选择是“不进入”；当 B 选{不进入，进入}时，A 的最优选择是“进入”。所以，有三个纳什均衡：{进入，（不进入，进入）}，{进入，（不进入，不进入）}和{不进入，（进入，进入）}，均衡的结果为：A 进入，B 不进入和 A 不进入，B 进入。“均衡结果”指 A 和 B 的均衡战略组合下的预测的 A 和 B 的行动组合，“均衡”指均衡战略的组合，两者是不同的概念。显然，在这里，多个不同均衡可以带来相同的均衡结果。细心的读者会发现：表 3.2 中用战略式表述的纳什均衡比图 3.1 中用扩展式表述的均衡多出一个，即{不进入，（进入，进入）}，其均衡结果为 A 不进入，B 进入。这个均衡实际上是不可信的，因为 B 威胁无论 A 是否进入他都要进入，若 A 相信 B 的威胁，给定 B 无论如何要进入，A 的最优选择当然就是不进入。但是，A 怎么能相信 B 的威胁呢？因为，只要 A 选择了“进入”，B 的最优选择就是“不进入”而不是“进入”。所以，在我们对局中人的理性行为假定下（而这正是博弈论的假定前提），应该预测 A 会选“进入”而 B 选“不进入”。所以，这个多出来的均衡应去掉。显然，当用战略式表述表达动态博弈时，可能将一些不可信的均衡包括进来，所以在动态博弈分析中常用的表述方式是扩展式而不是战略式表述。在这里，不可信的均衡中含有 B 对 A 的不可置信的威胁，应该从预测的均衡中去掉。我们将在第五章中对这个问题展开详细的讨论，并给出在较为一般的情况下去除这类甚至包括其他不可信均衡的方法，而剩下的均衡被称为“子博弈精炼纳什均衡”（sub–game Perfect Nash equilibrium）。类似地，也可用扩展式表述去表达静态博弈，两种表述方式是等价的。

我们在本章将专注于介绍局中人同时作出决策或者局中人在进行决策时不能观察到其他局中人的决策且不存在信息不对称的博弈模型，这类模型表示的博弈被称为“完全信息静态博弈”（perfect information static game）。

3.1 博弈模型的战略式表述

战略式表述的博弈模型需要在以下几个要素内容方面作出明确的确定，即局中人（player）——也称为参与人，指参与博弈的成员，可以是作为自然人的个人，也可以是企业、团体、组织机构、国家甚至国际联盟组织等。博弈论假定局中人是追求效用最大化的理性人。当局中人是企业、团体、组织机构甚至国家时，假定构成这类组织的自然人也是追求效用最大化的理性人。在具体的分析中，可以用利润最大化或其他目标函数来代替效用最大化，但这些不同的目标函数之间并不存在矛盾，它们都应被理解以效用最大化为一致性基础的相互可替代的表达方式。

记 n 为一个博弈中局中人的个数（$n=1,2,\cdots$），Γ 为所有局中人构成的集合，$\Gamma=(1,2,\cdots,n)$，i 为一个特定的局中人（$i=1,2,\cdots,n$），$i\in\Gamma$；

一个特别的局中人可能是"自然"，它是我们上述规定的例外，因为它既不是自然人也不是由自然人构成的组织，同时也不是追求效用最大化的行为主体，它往往表示一种博弈面临的环境或外生条件。

战略空间（strategy space）——每一个局中人可以选择的战略所构成的集合。一个战略（strategy）是指局中人选择行动的规则，而行动是指局中人的决策变量。一个战略告诉局中人在什么时候选择什么行动。譬如，毛泽东在制订中国与其他国家之间的军事战略时就通过十六个字简明扼要地概括出其主要思想："人不犯我，我不犯人，人若犯我，我必犯人"。在这个战略中，"犯"与"不犯"是两种不同的行动，该战略规定了什么时候选择"犯"，什么时候选择"不犯"。记第 i 个局中人的战略空间为 S_i；记 S_i 中的一个元素为 $s_i\in S_i$，$i\in\Gamma$；

信息——指局中人在博弈中拥有的相关知识，特别是有关其他局中人的特征和行动的知识。

支付函数——指局中人从博弈中获得的效用水平或利润水平或其他形式的目标函数。根据我们在前面给出的说明，无论是什么样形式的支付函数，它们一般都是以效用函数作为其基础的。

结果——这是一个内容较为广泛的概念，通常指研究者对博弈结束时所带来的各种感兴趣的效应或要素的集合，记第 i 个局中人的支付函数为 $u_i=u_i(s_1,\cdots,s_i,\cdots,s_n),i\in\Gamma$，局中人的支付不仅是该局中人自己所选战略的函数，而且还是所有其他局中人选择的战略的函数，这正是博弈论所强调的互动效应的数学描述。

均衡——所有局中人都选择的最优战略或行动的组合。在博弈论中，有不同的均衡概念，但基础性的均衡概念指的是"纳什均衡"。

我们一般将局中人、行动、结果统称为博弈规则，一个具体的博弈规则将决定相应的博弈均衡（但均衡不一定是唯一的）。

我们将一个博弈记为$G=\{S_1,\cdots,S_n;u_1,\cdots,u_n\}$，称这种表述方法为一个博弈的“战略式表述”（strategic form representation）。

我们在研究中经常会遇到一类较为简单的博弈，称为有限博弈，其定义如下：

定义 3.1　当局中人的个数n为有限数且每个局中人的战略空间中的元素只有限个时，称博弈为有限博弈（finite game）。

显然，我们在第二章中给出的“囚徒困境”博弈，“市场进入”博弈等都是有限博弈。

3.2 纳什均衡

在博弈论发展史上，均衡的概念有着一段发展演化的经历。从早期的占优战略均衡、重复剔除劣战略的占优战略均衡，直到后来的纳什均衡，博弈论的基本框架才告完成。继纳什之后，一些更为精致的均衡概念也被陆续提出，而且，这方面的研究还在不断深入。但是，纳什均衡概念无疑在博弈论发展史上占据有里程碑式的地位，它的提出标志着博弈论进入了一个有着完整方法论体系的新兴学科的迅猛发展时期。在介绍纳什均衡概念以前，我们先来回顾一下有关占优战略均衡和重复剔除劣战略的占优战略均衡的相关概念和理论。

3.2.1 占优战略均衡

所谓占优战略均衡，是指当所有局中人都选择各自的占优战略时所出现的战略组合。

为了表达上的简略，我们今后可将除第i个局中人所选择战略$s_i\in S_i$之外的其他所有局中人所选择的战略组合向量记为$s_{-i}=(s_1,\cdots,s_{i-1},s_{i+1},\cdots,s_n)$，$s_i\in S_i$，$i\in\Gamma$，又记$s=(s_1,\cdots s_{i-1},s_i,s_{i+1},\cdots,s_n)=(s_i,s_{-i})$。

定义 3.2　设$s_i,s_i^{'}\in S_i$，若满足

$$u_i(s_i,s_{-i})\le u_i(s_i^{'},s_{-i}),\quad \forall s_{-i}\in\prod_{j\ne i}S_j \tag{3.1}$$

其中$\prod_{j\ne i}S_j$表示$S_1,\cdots,S_{i-1},S_{i+1},\cdots S_n$构成的欧几里德乘积空间。则称$s_i$为$s_i^{'}$的劣战略，称$s_i^{'}$为$s_i$的占优战略。当不等式（3.1）对某些$s_{-i}$变成等式时，称$s_i$为$s_i^{'}$的弱劣战略，称$s_i^{'}$为$s_i$的弱占优战略。

当不等式（3.1）对任何s_{-i}都是严格不等式时，即不等式（3.2）成立。

$$u_i(s_i,s_{-i})< u_i(s_i^{'},s_{-i}),\quad \forall s_{-i}\in\prod_{j\ne i}S_j \tag{3.2}$$

则称s_i为$s_i^{'}$的严格或强劣战略，称$s_i^{'}$为s_i的严格或弱占优战略。

显然，$s_i^{'}$为s_i的占优战略就是在无论其他局中人选择什么战略情形下，局中人i选$s_i^{'}$都是相对于他选s_i为最优的战略。譬如，在前述给出的“囚徒困境”模型中，对于任一

犯罪嫌疑人，无论另一犯罪嫌疑人是选择“坦白”或是选择“抵赖”，他选择“坦白”都是相对于选“抵赖”为最优的战略。因此，“坦白”是“抵赖”的占优战略，而“抵赖”是“坦白”的劣战略，并且，“坦白”还是“抵赖”的强占优战略，“抵赖”是“坦白”的强劣战略。

定义 3.3　若 $s_i^{'}$ 是所有 S_i 中其他战略 s_i 的占优战略，则称 $s_i^{'}$ 为局中人 i 的占优战略，其中 s_i， $s_i^{'} \in S_i$， $i \in \Gamma$ 。

即是说当且仅当 $s_i^{'} \in S_i$（ $i \in \Gamma$ ）满足不等式（3.3）时， $s_i^{'}$ 是 i 的占优战略。

$$u_i(s_i, s_{-i}) \le u_i(s_i^{'}, s_{-i}) \qquad (3.3)$$

$$\forall s_i \in S_i，\ \forall s_{-i} \in \prod_{j \ne i} S_j$$

$$i, j \in \Gamma$$

当不等式（3.3）在某些 s_{-i} 处变为等式时，称 $s_i^{'}$ 为 i 的弱占优战略；当不等式（3.3）对任何 $s_{-i} \in \prod_{j \ne i} S_j$ 都是严格不等式时，称 $s_i^{'}$ 为严格或强占优战略。

但是，在博弈论中，我们通常所说的（弱）劣战略并不是指（弱）劣于任何其他战略的战略，只要某个战略（弱）劣于另一战略，我们就可称其为（弱）劣战略。

对于某局中人来说，假如存在某个强占优战略可选择，那么，我们就没有理由相信他会不选择这个战略。如果所有局中人都选择了各自的强占优战略。那么，我们可预测没有任何人可能改变已有的战略选择。从另一方面看，如果在一个博弈中，每一个局中人都存在一个强占优战略（显然，此时这种强占优战略必然是唯一的），我们可预测每一个局中人都会选择其各自的强占优战略。所以，我们就自然将这种可能出现的均衡状态当作一种可预测结果。于是，我们有如下定义：

定义 3.4　当一个博弈中的所有局中人都选择各自的强占优战略时，我们称博弈达到了占优战略均衡。

当所有局中人都选择了各自的强占优战略时，根据理性人假定，我们可预测任何一个局中人都不会改变其已有的战略选择，因而这种已有的战略组合是一种均衡。

需要指出，当我们把定义 3.4 中的“强占优战略”改为“占优战略”时，也会得到一种类似的均衡概念，在一些教科书中（譬如，张维迎 1996，P59），也将这种均衡称为“占优战略均衡”。但是，对于弱占优战略，我们不能预测局中人一定会选择它，因为还至少存在另一种可选择的战略，它给局中人带来的支付在其他局中人选择某些特定的战略组合时与该战略带来的支付相等。这样，我们不能预测博弈的最终战略组合选择是什么。或者说，在这种场合，可能出现的博弈结局有多种可能，因而存在多个不同的“占优战略均衡”。为了避免这种不确定性，博弈论中通常将“占优战略均衡 ”按定义 3.4 的方式给出其定义，即只考虑由强占优战略组合构成的“占优战略均衡”。我们在本书中也沿此惯例，在谈到“占优战略均衡”时，指的是按定义 3.4 给出的概念。

例 3.1　当我们在定义 3.4 中用“占优战略”取代“强占优战略”时，就可能存在多个“占优战略均衡”。表 3.3 中给出了一个例子。

表 3.3　存在四个“占优战略均衡”的博弈

		B		
		L	M	R
A	a	1, 1	2, 0	3, 1
	b	0, 2	1, 1	2, 2
	C	1, 3	2, 2	3, 3

当在定义 3.4 中用“占优战略”取代“强占优战略”时，A 有两个占优战略，即战略 a 和战略 c；B 也有两个占优战略，即战略 L 和战略 R；但 a、c 只是 A 的弱占优战略；L、R 只是 B 的弱占优战略。

显然，（c, L）、（c, R）、（a, L）和（a, R）都是在这种意义下的“占优战略均衡”。

我们在前章中给出的“囚徒困境”博弈就存在“占优战略均衡”，由于该博弈中每个局中人都有强占优战略“坦白”，所以，该博弈存在唯一的“占优战略均衡”即（坦白，坦白）。

3.2.2 重复剔除严格劣战略的占优战略均衡

尽管严格占优战略均衡或许是博弈模型能给出的一种很好的博弈结局预测，但存在严格占优战略均衡的博弈模型并不多见（存在占优战略均衡的博弈模型也很少），也就是说，运用严格占优战略均衡（甚至占优战略均衡）对博弈行为进行预测是不具普遍功用的，因为在许多博弈中，并不存在严格占优战略均衡（甚至占优战略均衡）。

然而，严格占优的概念对于寻找许多博弈模型的均衡解仍然具有启发性。如果我们将对“严格占优战略”的要求放松至“相对严格占优战略”，即不要求某战略比其他所有战略都严格为优，仅要求某战略只比另一战略严格为优，此时，我们有理由预测局中人不会选择另一战略（即劣战略），从而将其从博弈模型中剔除并得到一个新的博弈模型。假设其他局中人知道该局中人剔除了另一战略，而该局中人也知道其他局中人知道该局中人剔除了另一战略，博弈就在一个新的战略空间组合条件下进行。这种过程可以重复性地不断进行，因为剔除了某个劣战略的新博弈模型仍然可能含有新的属于其他局中人的劣战略，从而循着同样的方法将其剔除。当这种重复剔除劣战略的过程进行到最后时，倘若博弈模型只剩下唯一的一个战略组合，那么，可以预言这个战略组合是一种均衡，我们称其为“重复剔除劣战略的严格占优战略均衡”（iterated dominance equilibrium），简称为“重复剔除的占优战略均衡”。

定义 3.5　战略组合 $s^* = (s_1^*, \cdots, s_n^*)$ 称为重复剔除劣战略的严格占优战略均衡，如果

它是重复剔除严格劣战略后剩下的唯一的战略组合。当这种唯一的战略组合存在时，我们称该博弈是重复剔除占优可解的（dominance solvable）。

例 3.2　智猪博弈

在一个猪圈里养着一头大猪和一头小猪，在猪圈的一端放有一个猪食槽，在另一端安装有一个按钮，它控制着猪食的供应量。

假定：按一下按钮，就有 8 单位猪食进槽，但按动按钮需支出 2 单位成本；若大猪先到食槽，则大猪吃到 7 单位食物，而小猪仅能吃到 1 单位食物；若小猪先到，则大猪和小猪各吃到 4 单位食物；若两猪同时到，则大猪吃到 5 单位，小猪仅吃到 3 单位。

这里，每头猪都有两个战略：按或等待。见表 3.4 给出的战略式表述。

表 3.4　智猪博弈

		小猪	
		按	等待
大猪	按	3, 1	2, 4
	等待	7, –1	0, 0

显然，该博弈不存在占优战略均衡，因为尽管小猪有一个严格占优战略，但大猪并没有占优战略，所以，不能运用占优战略均衡概念求解此博弈；但我们却可用“重复剔除劣战略”的方法找出“重复剔除劣战略的占优战略均衡”解。

假定：小猪是理性的，因而她会选择严格占优战略“等待”，从而她会将劣战略“按”剔除。

假定大猪正确地预测到小猪会剔除“按”并选择“等待”，故而博弈就变成了表 3.5 给出的情形。

表 3.5　剔除劣战略后的智猪博弈

		小猪
		等待
大猪	按	2, 4
	等待	0, 0

显然，在新的博弈中，大猪有一个严格占优战略“按”，而此时“等待”仍是小猪的严格占优战略（小猪此时只有一种战略“等待”），假定大猪也是理性的（因而才被称为“智猪”），则大猪不会选择“等待”。因此，在这个新的博弈中，存在严格占优战略均衡（按，等待），即大猪选择“按”，小猪选择“等待”。

例 3.3　见表 3.6 给出的战略式博弈

表 3.6　一个战略式博弈的重复剔除劣战略过程

		B: L	B: M	B: R
A	U	1, 0	1, 2	0, 1
	D	0, 3	0, 1	2, 0

（a）

		B: L	B: M
A	U	1, 0	1, 2
	D	0, 3	0, 1

（b）

		B: L	B: M
A	U	1, 0	1, 2

（c）

该博弈不存在占优战略均衡，但局中人 B 有一个严格劣战略 R，因为战略 M 严格优于战略 R。假定 B 是理性的，则 B 将剔除战略 R；假定 A 知道 B 是理性的，因而 A 预测 B 会剔除 R，从而博弈变成表中（b）的情形；此时，A 有严格占优战略 U，假定 A 是理性的，他将剔除严格劣战略 D，从而博弈变成了表中的情形（c）；假定 B 知道 A 是理性的，并且 B 知道 A 知道 B 是理性的，从而 B 知道博弈变成了情形（c）中；此时，理性人 B 将选择 M，因而得到重复剔除劣战略的严格占优战略均衡（U，M）。在一个具体的博弈中，劣战略可能不是唯一的，即可能存在多个战略相对于其他战略来说是劣战略，此时，一个可能被提出的问题是：剔除劣战略的顺序（包括对同一局中人战略空间中的不同劣战略剔除顺序和不同局中人战略空间中劣战略剔除顺序）是否会影响到最后获得的均衡？不同的剔除劣战略顺序是否会给出不同的均衡？答案是这样的：对于有限博弈，倘若每一次剔除的是严格劣战略，则均衡与剔除的顺序无关；但对于无限博弈，剔除同一局中人的劣战略顺序可能与均衡有关，但在特定剔除方法假定下（指将同一局中人的所有劣战略一次性剔除的方法）均衡与不同局中人劣战略剔除顺序无关（见定理 3.1、推论 3.1.1 和注[1]、[2]）。于是我们有定理 3.1。

定理 3.1(蒲勇健，2001)　如果每一次剔除将同一局中人的所有劣战略剔除掉，则一

个战略式表述博弈在经过重复剔除（严格）劣战略后剩下的博弈与剔除的（局中人）顺序无关。

证明：设博弈的战略式表述为 $G=\{S_1^{'},\cdots,S_n^{'};u_1,\cdots,u_n\}$，我们只需证明经过剔除（严格）劣战略后剩下的博弈中各个局中人的战略空间与剔除的（局中人）顺序无关即可。

设 $A_i(1)$ 是由局中人 i 的所有劣战略构成的集合。

令 $S_i(0)=S_i$， $i=1,\cdots,n$ 。

$S_i(1)$ 是 S_i 中剔除所有 i 的劣战略后余下的集合，即 $S_i(1)=S_i\setminus A_i(1),i=1,\cdots,n$ 。

作新的战略式博弈 $G(1)=\{S_1(1),\cdots,S_n(1);u_1,\cdots,u_n\}$ （不妨设 $G(0)=G$ ）；

再设 $A_i(2)$ 是 $G(1)$ 中由局中人 i 的所有劣战略构成的集合，令 $S_i(2)=S_i(1)\ A_i(2)$，$i=1,\cdots,n$ 。

作新的战略式博弈 $G(2)=\{S_1(2),\cdots,S_n(2);u_1,\cdots,u_n\}$；

…………………………………………

一般地，设 $A_i(m)$ 为博弈 $G(m-1)=\{S_1(m-1),\cdots,S_n(m-1);u_1,\cdots,u_n\}$ 中由局中人 i 的所有劣战略构成的集合。

令 $S_i(m)=S_i(m-1)\setminus A_i(m)$， $i=1,\cdots,n$ 。

$G(m)=\{S_1(m-1),\cdots,S_n(m);u_1,\cdots,u_n\}$

显然有 $S_i(m)\subset S_i(m-1)$， $i=1,\cdots,n$；

令 $S_i^*=\bigcap_{m=0} S_i(m)$， $i=1,\cdots,n$ 。

今设有某一剔除过程，其中第 m 次剔除是对某个局中人的原战略空间中的劣战略进行的，而剔除后由该局中人原战略空间中余下的战略所构成的集合为 B_m $(m=1,\cdots)$ 。在这里，我们还需对剔除的方式作出一个特别的规定（尽管这种规定是特别的，但却是一种十分合理的假设），即假设每一次剔除都是同时将被剔除劣战略所属战略空间中所有的劣战略一并全部剔除[1]。

记该博弈经劣战略剔除后所得到的博弈为 $\tilde{S}^*=(\tilde{S}_1^*,\cdots,\tilde{S}_n^*;u_1,\cdots,u_n)$ 。

我们将证明 $\tilde{S}^*=S^*$（即 $\tilde{s}_i^*=s_i^*,i=1,\cdots,n$），由于 $\tilde{S}^*$ 是按任意（在上述规定的意义上）顺序剔除劣战略后所得到的博弈，故由此可证明定理为真。

我们先证明有 $S^*\subset\tilde{S}^*$，即 $S_i^*\subset\tilde{S}_i^*,i=1,\cdots,n$，使用数学归纳法：

显然，对于 B_1，必存在某个 i 使得 $B_1=S_i(1)$，设第二次进行劣战略剔除的战略空间为 S_j，显然有 $B_2\subset S_j$，而 $a_j\in S_j$ 为第二次剔除劣战略过程中被剔除的任一战略。我们在下面的证明过程中将不断重复使用这样一种逻辑，即当其他局中人的战略空间中的某些战略被剔除后，给定一个局中人的某个战略若在其他局中人的原有战略空间条件下是劣战略，则该战略在其他局中人的战略空间中的某些战略被剔除之后仍是劣战略，也就是说当其他局中人在较大的战略空间中选择战略时，a_j 为劣战略意味着当其他局中人在

较小战略空间中选择战略时 a_j 仍为劣战略。我们称该逻辑为 SSC 法则（Strategy Space Contract Law，战略空间收缩法则）。所以，$a_j \notin S_j(2)$，故

$B_2 \supset S_j(2)$

下面作归纳法假设：

设当 $m \geq 1$ 时有 $B_m \supset S_{j(m)}(m)$

其中 $j(k)$ 如此定义，即第 k 次进行劣战略剔除是对局中人 $j(k)$ 的战略空间中的劣战略进行的剔除，$k = 1,2,\cdots$。我们往证

$B_{m+1} \supset S_{j(m+1)}(m+1)$

设某 $a_j(m+1) \in S_{j(m+1)}$ 为第 $m+1$ 次剔除劣战略中被剔除的任一战略，据归纳法假设和基于 SSC 法则有

$a_{j(m+1)} \notin S_{j(m+1)}(m+1)$

于是有 $B_{m+1} \supset S_{j(m+1)}(m+1)$。

对于任意的 $i = 1,\cdots,n$，存在一个唯一的自然数序列集 $H_i = \{h_i\}$ 使得

$B_{h_i} \subset S_i$ 且 $B_{h_i^1} \subset B_{h_i^2}$ 当且仅当 $h_i^1 > h_i^2$，$h_i^1, h_i^2 \in H_i$

也就是说，第 h_i 次剔除劣战略是在局中人 i 的战略空间中完成的。

因前已证 $B_{h_i} \supset S_{j(h_i)}(h_i) = S_i(h_i)$，$h_i \in H_i$

故 $\bigcap_{h_i \in H_i} B_{h_i} \supset \bigcap_{h_i \in H_i} S_i(h_i)$

因 $\bigcap_{m=0} S_i(m) \subset \bigcap_{h_i \in H_i} S_i(h_i)$

得 $\bigcap_{h_i \in H_i} B_{h_i} \supset \bigcap_{m=0} S_i(m) = S_i^*$

据 B_m 和 $\tilde{S}_i^*$ 的定义显然有 $\tilde{S}_i^* = \bigcap_{h_i \in H_i} B_{h_i}, i = 1,\cdots,n$

故　$\tilde{S}_i^* \supset S_i^*$，$i = 1,\cdots,n$

即　$\tilde{S}^* \supset S^*$

下面我们再证明反过来也有　$\tilde{S}_i^* \subset S_i^*$。

显然，按照我们规定的剔除方式，$B_1 = S_{j(1)}(1)$，$B_2 = S_{j(2)}(1)$，有可能某些 $S_i(k)$ 在 $G(k)$ 中没有劣战略，$k = 0,1,\cdots$，因而对任何 B_m 都有 $B_m \not\subset S_i(k), m = 1,\cdots$。但为了以下数学讨论上的方便，我们可根据需要在 $\{B_m\}$ 中添加某些 B_j 使 $B_j \subset S_i(k)$，$k = 0,1,\cdots$；但实际上对应于这些 B_j 的剔除并没有发生。显然，经过这样调整后的剔除实际上将与调整前的剔除过程具有同样的剔除结果。这样，我们不妨假设所讨论的 $\{B_m\}$ 已经经过如此调整。按此规定，我们可定义第 h_{i1} 次剔除是第一次对局中人 i 的战略空间中的劣战略进行剔除，据 SSC 法则有

$B_{h_{i1}} = S_i(1)$　（3.4）

设 $H(1) = \max\{h_{i1}, i = 1,\cdots,n\}$

于是 $B_m \subset S_{j(m)}(1)$，当 $m > H(1)$

定义 $H(2)=\max\{h_{i1}, h=1,\cdots,n\}$，其中 h_{i2} 的定义为：第 h_{i2} 次剔除是第 $H(1)$ 次剔除后第一次对局中人 i 的战略空间中的劣战略进行的剔除（若原来的 $\{B_m\}$ 中没有这种剔除，我们在这里不妨按前述思路在 $\{B_m\}$ 中添加一次对 S_i 中的劣战略"剔除"，但实际上并未发生真实的"剔除"）。

由式（3.4）和 SSC 法则，有

$B_{h_{i2}} \subset S_i(2)$，$i=1,\cdots,n$

下面，我们用归纳法证明一般有

$$B_{h_{im}} \subset S_i(m), \quad i=1,\cdots,n, m=1,\cdots \tag{3.5}$$

其中 h_{im} 是如此定义的：第 h_{im} 次剔除是在第 $H(m-1)$ 次剔除后第一次对局中人 i 的战略空间中的劣战略进行的剔除（若原来的 $\{B_m\}$ 中没有这种剔除，我们在这里不妨按前述思路在 $\{B_m\}$ 中添加一次对 S_i 中的劣战略"剔除"，但实际上并未发生真实的"剔除"），而 $H(m-1)$ 的定义用归纳法定义。已定义了 $m=2,3$ 时的 $H(m-1)$，假定 $H(m-2)$ 已有定义，则

$H(m-1)=\max\{h_{im-1}, i=1,\cdots,n\}$

其中，h_{im-1} 的定义为：第 h_{im-1} 次剔除是在第 $H(m-2)$ 次剔除后第一次对局中人 i 的战略空间中的劣战略进行的剔除（若原来的 $\{B_m\}$ 中没有这种剔除，我们按前述方式不妨在 $\{B_m\}$ 中添加一次对 S_i 中劣战略的"剔除"）。

令 $H(m)=\max\{h_{im}, i=1,\cdots,n\}$

其中 h_{im} 的定义为：第 h_{im} 次剔除是在第 $H(m-1)$ 次剔除后第一次对局中人 i 的战略空间中的劣战略进行的剔除（若原来的 $\{B_m\}$ 中没有这种剔除，我们按前述方式不妨在 $\{B_m\}$ 中添加一次对 S_i 中劣战略的"剔除"）。

再令 $H(m+1)=\max\{h_{im+1}, i=1,\cdots,n\}$

其中 h_{im+1} 的定义为：第 h_{im+1} 次剔除是在第 $H(m)$ 次剔除后第一次对局中人 i 的战略空间中的劣战略进行的剔除（若原来的 $\{B_m\}$ 中没有这种剔除，我们按前述方式不妨在 $\{B_m\}$ 中添加一次对 S_i 中劣战略的"剔除"）。

据式（3.5）和 SSC 法则，有

$B_{h_{im+1}} \subset S_{i(m+1)}, i=1,\cdots,n$

按我们的前述假设，对任何 $m=1,\cdots$，h_{im} 都存在，且 $B_{h_{im+1}} \subset B_{h_{im}}$ 和 $h_{im+1} > h_{im}$，$m=1,\cdots$。

任给 $a_i \in \bigcap_{m=1} B_{h_{im}}$，对任意的 $B_h \subset S_i$，必有某 $B_{h_{im}} \subset B_h$，只要 $h_{im} > h$，故

$a_i \in B_h$，所以 $a_i \in \tilde{S}_i^*$

反之，任给 $a_i \in \tilde{S}_i^*$，必有 $a_i \in B_{h_{im}}$，$m=1,\cdots$

于是 $a_i \in \bigcap_{m=1} B_{h_{im}}$

得 $\bigcap_{m=1} B_{h_{im}} = \tilde{S}_i^*$

由式（3.5）还有

$\bigcap_{m=1} B_{h_{im}} \subset \bigcap_{m=1} S_i(m) = S_i^*$

所以得到 $\tilde{S}_i^* \subset S_i^*$， $i = 1, \cdots, n$

即 $\tilde{S}^* \subset S^*$

因而得到 $\tilde{S}^* = S^*$

证毕！[2]

推论 3.1.1 若博弈是重复剔除占优可解的，则重复剔除的占优战略均衡与剔除劣战略的（局中人）顺序无关。

证明：当博弈是重复剔除占优可解的，则 $\tilde{S}_i^*$ 都是单一元素集合，即 $\tilde{S}_i^*$ 中只有一个元素 a_i^*， $i = 1, \cdots, n$ ，均衡为 $\tilde{S}^* = (a_1^*, \cdots, a_n^*)$ ，据定理 3.1 知 $\tilde{S}^*$ 与剔除（局中人）顺序无关。

证毕。

在定理 3.1 和推论 3.1.1 中，我们都假定了剔除的是严格劣战略。一个十分自然的问题是，倘若被剔除的包括有非严格劣战略，剔除的顺序与均衡结果有关吗？答案是肯定的。当被剔除的战略中含有非严格劣战略时，剔除顺序不同则剩下的博弈是不同的，当剩下的博弈恰好就是均衡战略组合时，剔除的顺序不同则剩下的博弈均衡也可能不同。在表 3.7 中，我们给出这样一个例子。

表 3.7　剔除非严格劣战略将导致均衡结果与剔除顺序有关的例子

		B: L	B: M	B: R
A	U	1, 3	2, 1	2, 3
A	H	0, 3	0, 1	0, 2
A	D	0, 2	0, 1	0, 4

当我们以顺序 D、R、M、H 剔除时，均衡为（U, L）；但若剔除顺序改为 M、H、L、D，则均衡为（U, R）[3]。当我们通过剔除严格劣战略方式对表 3.7 中的战略式博弈进行重复剔除劣战略，将发现它不是重复剔除占优可解的。

3.2.3 关于重复剔除的占优战略均衡的几点评论

尽管重复剔除劣战略方法在求解博弈模型均衡解上比起占优战略均衡方法前进了一步，但这种方法仍然存在严重不足。除了许多博弈模型不是重复剔除占优可解，从而

这种方法不具普遍适用性外，一个致命的缺陷是这种方法对局中人过高的理性要求实际上超出了现实中大多数自然人的真实理性程度。譬如，在例 3.3 中，我们要通过重复剔除劣战略方法得到均衡（U, M），就需要假定不仅 A 和 B 都是理性的，而且还要假定 A 知道 B 是理性的，还要假定 B 知道 A 是理性的并且 B 知道 A 知道 B 是理性的。一般地，A 是理性的和 B 是理性的并不意味着 A 知道 B 是理性的和 B 知道 A 是理性的，更不意味着 B 知道 A 知道 B 是理性的和 A 知道 B 知道 A 是理性的。因此，假定 A 知道 B 是理性的实际上对 A 的理性程度作了比 A 是理性的和 B 是理性的假定更高的理性程度假设；假定 A 知道 B 知道 A 是理性的又更进一步对 A 的理性程度作了更高的假设。显然，当博弈中的局中人战略空间增大时，剔除劣战略的次数会增加，为了得出重复剔除占优均衡，就需要对局中人的理性程度作出更高的假设，如假定 A 是理性的、B 是理性的、A 知道 B 是理性的、B 知道 A 是理性的、A 知道 B 知道 A 是理性的、B 知道 A 知道 B 是理性的、A 知道 B 知道 A 知道 B 知道 A 是理性的……等等。在一般的博弈场合，我们若对所有局中人的理性程度作出最高程度的假设，即假设所有局中人是理性的、所有局中人知道所有局中人是理性的、所有局中人知道所有局中人知道所有局中人是理性的，……，则这种假设在博弈论中被称为“理性是局中人的共同知识（Common knowledge）”假设。

在博弈论中，许多模型在“理性是局中人的共同知识”假定下可轻易地求得均衡解，但是，模型解往往与现实中的真实博弈结果大相径庭。造成这种矛盾的原因是多方面的，但“理性是局中人的共同知识”假定对局中人理性程度作出过高假设也许是主要的原因。譬如，在中国古典名著《三国演义》中，曹操在赤壁大战后败走华容道。当时，曹操败军在踏上华容道之前曾面临两条撤退道路可选择，一是华容道，另外还有一条小路可逃。孔明事先令关羽率精兵在华容道上设下埋伏，等待曹军败走此路时活捉曹操。被誉为旷世奇才的孔明军师令关羽在华容道上点燃草堆发出缕缕白烟。当曹操败至分叉路口时，前兵来报：前有两条路可走，但华容道上有缕缕轻烟浮动，料有伏兵。众将闻此言，皆云选另一条道避开华容道逃命。曹操催马来到队前，举手加额，燎望远处华容道上的轻烟，思想良久，踌躇几分，忽转过身来面向众将抚掌大笑，遂令大军奔华容道而去。众将皆莫名其妙，皆云：“烽烟起处，必有军马，何故反走这条路？”操曰：“岂不闻兵书有云：虚则实之，实则虚之。孔明多谋，以为我看见烟火就会避开华容道走另一条路，而他却正好派兵在此道上等着我，我才不上他的当呢！”于是，曹操令大军沿华容道前进，但行进不久就望见关羽横刀勒马于华容道上……[4]。显然，曹操在理性程度上与孔明相比是稍逊一筹的。孔明和曹操都是理性的，孔明知道曹操是理性的，曹操也知道孔明是理性的（故他知道：“孔明多谋”）。因此，孔明知道曹操看见华容道上的烟火后会走另一条路，曹操知道孔明的诡计，因此猜想孔明会在此路上设伏兵，故而曹操以为识破了孔明计谋而令大军直奔华容道；但孔明知道曹操知道孔明知道曹操看见华容道上的

烟火后会走另一条路，故孔明知道曹操会如此所想而直冲华容道来，所以孔明令关羽在华容道上等着曹操，结果将曹操逮个正着。[5]这个故事的真实性令人怀疑，因为倘若曹操也知道孔明知道曹操知道孔明知道曹操看见华容道上的烟火后会走另一条路，则曹操会避开华容道而走上另一条路；但故事到此还未说完，如果孔明又知道曹操也知道孔明知道曹操知道孔明知道曹操看见华容道上的烟火后会走另一条路，则孔明又会令关羽在另一条道上堵截曹操而不是伏兵于华容道，而这又正好将曹操逮个正着。但故事仍未完结，如果……，我们可以想象，这种对孔明和曹操理性程度的假设可以无限增高下去，而结果并不能明确告诉我们到底关羽能否逮住曹操。我们将在后面引入了混合战略博弈后，再来重新将这个故事表述为博弈问题，并证明只要我们不假定孔明在理性程度上高出曹操一筹，则孔明不一定能令关羽活捉住曹操，还要证明当我们假设孔明和曹操在理性程度上不分伯仲时（假设他俩的理性程度相同），则曹操有一半的机会逃脱。

在中国古典文献中，我们还可以找到另一个类型的例子。据《韩非子》所述，战国时孙武的后代孙膑因受同窗庞娟的迫害致残而远走他乡，最后流落在齐国名将田忌帐下作谋士。一日孙膑见田忌闷闷不乐，便问田忌有何不快之事。田忌告诉孙膑，近来齐威王（齐国国王）常要他与其赛马，赛马规则是每次双方各出三匹马，一对一比赛三场，每一场的输方要赔一千斤铜给赢方。齐威王的三匹马和田忌的三匹马按实力都可分为上、中、下三等，而齐威王的上、中、下三匹马都分别比田忌的上、中、下三匹马略胜一筹，因此田忌每次都是连输三场，要输掉三千斤铜。孙膑闻得此事，心中筹划如何为田忌献计赢得赛马。结果，经调查研究发现，田忌的上马虽不如齐威王的上马，却要比齐威王的中马和下马都要好，同样，田忌的中马则比齐威王的下马要好一些。于是，孙膑为田忌出奇计：先用田忌的下马对抗齐威王的上马，再用田忌的上马对抗齐威王的中马，最后才用田忌的中马对抗齐威王的下马。这样，田忌可以输掉第一场作为代价而赢得后二场比赛，每次比赛可净赢齐威王一千斤铜。这个故事告诉我们在实际生活中有可能处于绝对劣势的一方依靠计谋仍可战胜强者。但是，在这个故事中，我们实际上假设了齐威王是“傻子”——齐威王凭什么要用自己的上马对田忌的下马呢！当田忌出下马时，齐威王完全可以出下马取胜，当田忌出上马时，齐威王出上马可胜，当田忌出中马时，齐威王可出中马胜之；但反过来，当齐威王出下马时，田忌又可换出中马取胜，当齐威王出上马时，田忌可出下马输一场，然后当齐威王出中马时，田忌再出上马赢得最后一场，从而净赢一场；但又反过来，若田忌开始换出中马，齐威王则可换出中马，……。我们看到，这个游戏完全类似于前面提到的《三国演义》中华容道上的关羽与曹操的故事。在开始，田忌是“傻子”，他的理性程度稍逊于齐威王，故而傻乎乎地用自己的上马对齐威王的上马，用自己的中马对齐威王的中马、用自己的下马对齐威王的下马。然后，经孙膑点拨后，田忌的理性程度提高了，但这时齐威王又成了“傻子”，其理性程度反过来逊于田忌（和孙膑），因为他居然在拥有绝对优势条件下输掉一千斤铜——他

居然愚蠢地用自己的上马对田忌的下马、用自己的中马对田忌的上马、用自己的下马对田忌的中马。

如果这两个历史故事是曾经真实发生过的，那么，现实中的博弈者或许并不满足重复剔除劣战略中所假定的理性条件，理性也更不是局中人的共同知识。譬如，曹操在华容道上被关羽捉住，就是因为曹操的理性程度与孔明相比略低一个层次，因为尽管曹操知道孔明是理性的，故曹操以为孔明通过令关羽在华容道上施放烟火来诱骗其上当走上另一条路而此路上正好设有伏兵，所以曹操径直踏上华容道；但是，孔明不仅知道曹操是理性的，而且还知道曹操知道孔明是理性的，故而令关羽在华容道上等着曹操，结果将曹操逮个正着。在“田忌赛马”故事中，开始田忌居然不是理性的；因为在给定齐威王的出马顺序下他未能最优地设计自己的出马顺序（给定齐威王的出马顺序为上、中、下，田忌的最优出马顺序为下、上、中，但田忌居然按上、中、下顺序派出赛马）。反过来，给定田忌的出马顺序为上、中、下，齐威王按上、中、下出马是最优的，所以齐威王是理性的。

尽管如此，博弈论并不假设不同的局中人具有不同的理性程度，当我们假定曹操比孔明“傻”或齐威王比孙膑和田忌“蠢”时，曹操被孔明智胜或齐威王输给孙膑与田忌是理所当然的。博弈论要研究的是在局中人的理性程度完全相同的条件下，局中人如何在战略选择下达到某种均衡。我们将在后面证明，当假定曹操与孔明具有同样的理性程度或者田忌与齐威王同样聪明时，博弈必是一种被称之为“混合战略博弈”的博弈过程，其中局中人以随机方式选取战略，且“平均看来”曹操有一半的机会逃脱而田忌在某种均衡中仍将输给齐威王。所以，“田忌赛马”故事实际上与博弈论没有什么关系，该故事属于决策论范畴。

有一些介绍博弈论的书将“game theory”译为“对策论”，这种译法不太恰当，因为“对策论”易使人将其与人们常说的“什么什么对策”相混淆，如“地区发展对策”，“犯罪整治对策”，“铁路春运期间安全对策”等等，而这些问题实际上属于“决策论”所研讨的范围，但“决策论”与博弈论是存在很大区别的。决策论是研究局中人在给定其他环境参数条件下的最优选择问题，如孙膑在给定齐威王出马顺序为上、中、下时为田忌筹划了按“下、上、中”顺序出马的最优选择。博弈论研究的是当局中人充分考虑到其他局中人对其战略选择的反应后（即局中人都具有同样充分的理性时）进行最优战略的选择。“田忌赛马”故事中孙膑和田忌未能考虑齐威王对他俩的选择会有什么反应，所以属于决策论问题。当我们在后面将这个故事改为博弈论问题时，齐威王“平均”看来仍会赢得赛马，因为他拥有绝对优势。

为了看清楚博弈论与决策论的区别，我们在下面再给出一个例子。一般地，在决策论中，如果是单人决策，当对应一个给定选择的所有可能状态下的支付同时下降时，则当事人的效用水平绝不会提高，但对于博弈论来说，结论却恰恰可能相反。在表 3.8 中，

我们给出一个战略式博弈。

表 3.8　作为决策论的一个反例的博弈

		B	
		L	R
A	U	1, 3	4, 1
	D	0, 2	3, 4

显然，此博弈的重复剔除占优均衡为（U, L），A 和 B 的支付分别是 1 和 3。现假设给定 A 选 U，B 选 L 和 R 时 A 的支付同时下降 2 个单位，从而得到如表 3.9 给出的一个新博弈。

表 3.9　将表 3.8 中的博弈支付调整后的新博弈

		B	
		L	R
A	U	–1, 3	2, 1
	D	0, 2	3, 4

根据决策论，这种调整不会使 A 获益。当给定 B 的选择为 L 时，无论 A 选什么，其效用会下降。但在博弈论里，若这种变化为 B 所知，则 A 的效用会上升，因此时 B 知 D 是 A 的占优战略，B 就会改选 R 而不是 L，A 就获得 3 单位效用而不是原来的 1 单位效用。类似情况也出现在当一个局中人的选择空间缩小或信息质量下降的时候。在单人决策里，选择的空间缩小或信息质量的下降绝不可能是件好事，但在博弈中，却可能使当事人受益。我们将在以后的例子中看到这一点。

在博弈论中，与“理性是局中人的共同知识”假定发生有趣关联的还有一个博弈故事，称为电子邮件博弈（E–mail game）。该故事说的是在欧洲大陆还处于巧取豪夺、强盗横行和诸侯割据的中世纪，有三个地方军阀各自占据了一个城堡。我们用 A、B 和 C 分别代表这三个军阀。有一天，军阀 A 获得重要情报，即发现军阀 C 拥有巨量财宝。假定军阀 A 和 B 联合起来进攻军阀 C 可保证有足够的力量攻破 C 的城堡并打败军阀 C，但军阀 A 和 B 中任何一个单独进攻 C 都会招致失败并被 C 击溃。现在，军阀 A 写一封信给军阀 B，信中告诉 B 他所获得的情报，并建议 B 与 A 联合起来于第二天向 C 的城堡发起进攻，共同瓜分从 C 那里抢夺来的财宝。假定 C 的城堡中藏匿着的财宝数量巨大，足以补偿两支进攻军队在战斗中蒙受的损失。于是，A 将写好的信封好后令一位士兵向 B 送去。我们现在的问题是：尽管 A 和 B 组成联军向 C 发起进攻是一个帕累托最

优（对 A 和 B 来说）的选择，但 B 收到 A 的信后果真会于第二天派出军队向 C 进攻吗？乍看起来，答案好象是“B 会于第二天向 C 进攻”，但且慢，如果我们沿用博弈论中的理性人思维逻辑进行推演，所获得的结论却恰好相反，即“B 不会于第二天向 C 进攻”。这是因为，尽管 B 收到 A 的信并读了信中内容，但 B 不能保证 A 知道 B 收到 A 的信，从而 B 不能保证 A 会在第二天也向 C 进攻。如果 A 认为 B 未收到信，则 A 不敢贸然向 C 进攻（根据假设，A 和 B 中任何一方单独向 C 进攻都必然招致毁灭性打击），如果 A 不向 C 进攻，B 也不敢单独向 C 进攻。所以，B 又需要给 A 回一封信，告诉 A 他已收到 A 的来信并赞同 A 在信中所提的建议，于是令那位送信士兵将回信带回去交给 A。现在的问题是，当 A 收到 B 的回信后，A 会在第二天向 C 进攻吗？答案仍然是“不会”。因为 A 不知道 B 是否知道 A 收到 B 的回信，因而 A 不能保证 B 会向 C 进攻。如果 B 认为 A 可能未收到 B 给 A 的回信，那么，B 就不敢贸然向 C 进攻。B 不向 C 进攻，A 也就不敢向 C 进攻。这样，A 还需再向 B 写一封回信，说明已收到 B 的回信并于第二天向 C 进攻。显然，当 B 收到 A 的回信后，B 仍然不敢向 C 进攻，因为 B 不能保证 A 知道 B 收到了 A 的回信。于是，还需 B 再向 A 写第二封回信，……。于是，只要这种信件来回的次数是有限的，A 和 B 组成联军向 C 发起进攻的事就不会发生。我们可以用逆推法来说明其中的道理：不妨设经过有限次信件来回后，最后一封信传到 B 手中，此时 B 仍不敢向 C 进攻，因为 B 不知道 A 是否知道 B 已收到最后一封信，如果 B 未收到，A 就不能保证 B 是否知道 A 已收到 B 给 A 的倒数第二封信，倘若 B 认为 A 未收到倒数第二封信，B 就不能保证 A 是否知道 B 已收到 A 给 B 的倒数第三封信，……，沿此思路，我们不难得出：B 不能保证 A 是否知道 A 给 B 的第一封信，从而 A 可能认为 B 未收到第一封信，这样，A 就不敢进攻 C，给定 A 不进攻 C，B 也不会进攻 C。显然，若最后一封信传在 A 手中，同样的分析将导致相同的结论。也许我们可以设计一种机制来使得这种没完没了的信件传递经有限次后中止，并实现 A 和 B 向 C 的联合进攻。譬如，A 可在第 5 封信上写下这样的字句：当你收到我的这封信（第 5 封信）后，你就不要给我回信了，并于明天向 C 发起进攻。显然，即使这样，A 仍然不会向 C 进攻。这是因为，如果 B 不向 A 回第六封信，A 怎么能保证 B 会收到他向 B 写的第 5 封信呢？如果 B 未收到第 5 封信，B 就可能认为 A 未收到 B 向 A 写的第 4 封信；如果 A 未收到第 4 封信，A 就可能认为 B 未收到 A 给 B 写的第 3 封信；如果 B 未收到第 3 封信，B 就会认为 A 未收到第二封信；如果 A 未收到第二封信，A 就认为 B 可能未收到他给 B 写的第一封信；如果 B 未收到第一封信，B 就不会进攻 C，给定 B 不会进攻 C，A 也不会进攻 C，而 A 不进攻 C 又导致 B 会作出不进攻 C 的决定。显然，我们要使 A 和 B 联合进攻 C 这件事干成，只有作出“理性是局中人的共同知识”的假定，即假定“A 知道 B 知道 A 知道 B 知道 A 知道 B 知道……会向 C 进攻”，或者假设 A 与 B 之间来往无限次信件，但这在实际生活中是不可能的，即使是传输速度极快的电子邮件或无线电话，

严格按上述逻辑也仍然会导出同样结论，因为上面对电子邮件博弈的分析只要在信息传递的速度是有限的条件下都是成立的。如果我们不相信有超距作用，那么，电子邮件博弈的逻辑总是成立的。

电子邮件博弈的本质是局中人之间难以在多个博弈均衡之间进行选择，因为任何一个局中人都不能预测其他局中人会选择哪一种均衡战略。在表 3.10 中，我们给出电子邮件博弈的一种战略式表述。

表 3.10　电子邮件博弈，其中 a 是充分大的正数

		B	
		进攻	不进攻
A	进攻	1, 1	–a, 0
	不进攻	0, –a	0, 0

在表 3.10 中，有两个纳什均衡（进攻，进攻）和（不进攻，不进攻）。A 的两个战略“进攻”和“不进攻”都是 A 的均衡战略，B 的两个战略“进攻”和“不进攻”也都是 B 的均衡战略。但是，如果 A 不能预测 B 会选择“进攻”还是“不进攻”，则 A 就不能决定是否选择“进攻”。我们假设局中人 A、B 将失败或被消灭视为极为可怕的事情，则表 3.10 中的 a 就是一个充分大的正数。A 充分大意味着当 A 不能肯定 B 一定会选“进攻”时（从而 A 预测 B 有可能选择“不进攻”），A 就一定会选择“不进攻”。同样，B 也会在不能肯定 A 一定会选“进攻”下选择“不进攻”。在这个博弈中，均衡（进攻，进攻）是帕累托最优的，但真正出现的均衡都是（不进攻，不进攻）。假若 A 与 B 之间能进行沟通，似乎可以合谋达成帕累托最优均衡（进攻，进攻），但前述分析表明这种合谋又是不能达成的。一般地，博弈中常出现多重均衡，局中人之间在沟通上的局限和不完全性使得局中人在均衡战略选择上难以达成帕累托最优均衡，除非我们假定局中人之间可以非现实地进行无数次沟通。

另外，由于重复剔除的占优战略均衡对局中人的理性程度要求过高，不允许局中人犯丝毫错误，因而据此均衡对实际结果进行预测往往欠合理。表 3.11 给出一个战略式博弈，其中（U, L）是重复剔除占优均衡，但如果我们做一个实验，可能会有很多人会选 D 而不是 U。因为尽管如果局中人 A 百分之百地相信 B 是理性的因而会选 L，从而 U 是 A 的最优选择，但是，如果 B 有千分之一的概率选 R，A 选 D 就优于 U。所以，若现实中的 A 和 B 不是重复剔除的占优战略均衡中所要求的那样具有绝对理性，则（D, L）就可能是一个较（U, L）更好的预测。

表 3.11　重复剔除的占优战略均衡作为一个不好的预测的例子

		B	
		L	R
A	U	8, 10	–1000, 9
	D	7, 6	6, 5

3.2.4 存在重复剔除的占优战略均衡博弈模型的几个应用例子

我们曾在 2.1.3 小节中提到这样一种思路，即通过博弈论模型可以使人们认识到许多在表面上看来毫无共同之处的经济现象（甚至包括社会、军事、生物学在内的其他现象）之间的相同的本质。除了在那里所给出的“囚徒困境”模型外，我们在这里再看看其他一些模型能对什么样的经济现象作出解说。在这里，我们将视野局限于一些存在重复剔除的占优战略均衡的模型。

例 3.4　用智猪博弈模型解释股市现象

智猪模型在例 3.2 中已给出，其重复剔除的占优战略均衡为{大猪选择“按”，小猪选择“等待”}。在股市中，大户是“大猪”，散户是“小猪”。大户要自己搜集信息，进行投资分析，而散户只是选择“跟大户”。

例 3.5　用智猪博弈模型解释股份公司治理结构

在股份公司中，大股东是“大猪”，小股东是“小猪”。尽管股东承担着监督经理的职能，但大股东和小股东从监督中得到的收益并不一样。监督经理需要付出搜集信息和花费时间的成本。在监督成本相同的情况下，大股东从监督中得到的好处显然多于小股东。因此，均衡状态下大股东担当起搜集信息、监督经理的责任，而小股东则搭大股东的便车。正是基于这种理由，大股东有选举投票权，通过投票权行使对经理的监督。

例 3.6　用智猪博弈模型解释广告市场现象

在广告市场上，花钱制作广告的多数是大企业，小企业做广告是得不偿失的。小企业通常模仿大企业的产品，等待大企业通过广告打开市场后出售廉价产品。

例 3.7　用智猪博弈模型解释企业技术创新

大企业进行研究开发、花钱进行技术创新，而小企业等待大企业开发出新技术新产品后通过模仿或创新成果的扩散生产并出售类似产品。

例 3.8　用智猪博弈模型解释公共品的提供

在一些乡村，通常由有钱的富户出钱修路，因为富户家中来往的宾客多，其出门次数多，且无论是他们自己还是宾客都多半使用高档交通工具如小轿车，因为富户从修路中获得的好处较多；穷人则穿着破鞋走路，路的好坏对他们的效用影响不大。

例 3.9　用智猪博弈模型解释改革中的积极性差异

在经济或社会体制改革中，我们常看到某些人或社会阶层十分积极地推动改革，而

其他人或社会阶层则相对冷漠。这是因为同样的改革带给一部分人的好处可能比另一部分人大得多。此时，前一部分人比后一部分人更有积极性改革，改革往往就是由这些“大猪”推动的。譬如，年轻人可能在改革中获益较大，因而年轻人一般较为积极地投入改革。

通过重复剔除劣战略来获得均衡的方法还可用于存在连续性战略空间的博弈，例 3.10 给出一种示例。

例 3.10　古诺寡头竞争模型：重复剔除的占优战略均衡

数理经济学的先驱古诺（Cournot，1838）早年提出的寡头竞争模型可以说已经含有了纳什均衡的思想了，它比纳什（1950）给出的定义提早了 100 多年。该模型的特点是局中人在选择战略和行动时是以其他局中人的战略选择给定情况下进行的。如果给定其他局中人的选择，某局中人现在选择的战略是最优的；反过来，如果给定该局中人的已选择战略，其他局中人已选定的战略仍是最优的，当该局中人是任意一位局中人时该结论都成立，则古诺定义这是一种均衡。显然，古诺在这里给出的均衡就是纳什均衡[6]。在古诺寡头竞争模型中，局中人是两个生产同质产品并在同一市场上展开竞争的企业，记为企业 1 和企业 2。每个局中人的战略空间有两种表达，通常用企业可选择的产量作为战略空间元素，但也可将企业选择的价格作为其战略空间元素。在这里，我们将局中人的战略空间定义为企业可选择的产量范围，而局中人的支付为其可获利润。于是，我们有如下战略式表述的博弈：

$S_1 = S_2 = [0,\infty)$，$C_i(q_i)$ 为企业 i 的成本函数，q_i 为企业 i 的产量，$i = 1,2$；设逆需求函数为 $P = P(q_1 + q_2)$，其中 P 为产品价格。

局中人 i 的支付函数为

$$\pi_i(q_1, q_2) = q_i P(q_1 + q_2) - C_i(q_i), \qquad i = 1,2$$

对于局中人 i，给定局中人 $j(j \neq i)$ 的产量 q_j，其最优产量 q_i（利润最大化产量）满足以下一阶条件：

$$\frac{\partial \pi_i}{\partial q_i} = P(q_1 + q_2) + q_i P'(q_1 + q_2) - C'_i(q_i) = 0, \qquad i = 1,2$$

该式定义了局中人 i 对局中人 $j(j \neq i)$ 的反应函数：

$$q_i^* = R_i(q_j), \qquad i \neq j, i, j = 1,2$$

令 $q_i^m = R_i(0)$ 为局中人 i 的最优垄断产量（即局中人 j 不生产时局中人 i 的最优产量）。为了简化数学过程，假定两个局中人具有相同的常数平均成本，即 $C_i(q_i) = Cq_i, i = 1,2$，且逆需求函数取如下线性形式：

$$P = a - (q_1 + q_2)$$

于是，一阶条件变为

$$a - (q_1 + q_2) - q_i - c = 0, \qquad i = 1,2$$

我们定义“反应函数”为

$q_i^* = \frac{1}{2}(a - q_j - c)$，$i \neq j$，$i, j = 1,2$

在图 3.3 中，我们描出了两条反应函数的图象。

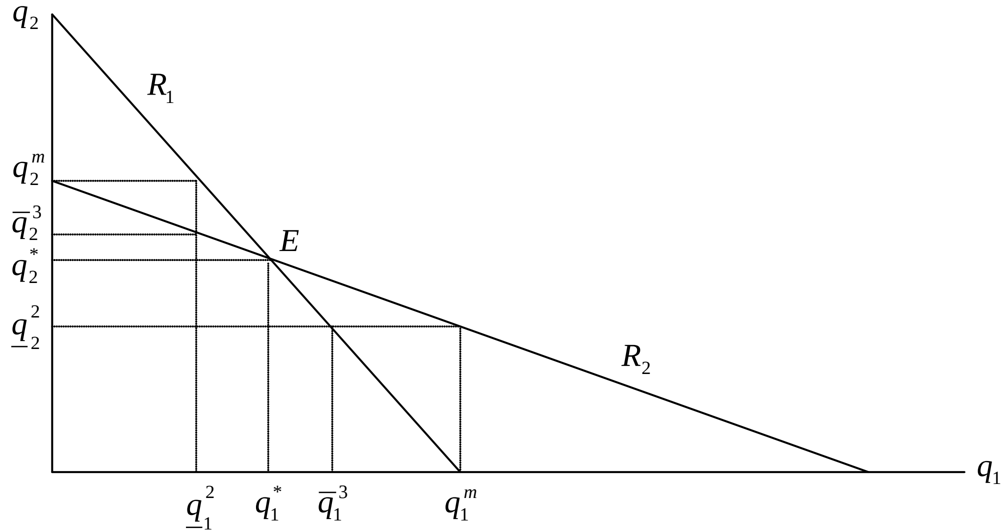


图 3.3　古诺模型：重复剔除过程

从图中看出，$q_i \in (q_i^m, \infty)$ 是劣战略，因没有企业会选择大于垄断产量的产量，在图 3.3 中，$q_i \in (q_i^m, \infty)$ 是任何情况下局中人 i 也不会选择的产量，故第一轮剔除劣战略后得到 $S_i(1) = [0, q_i^m]$，$i = 1,2$。然后，给定局中人 i 知道局中人 j 不会选择 $q_j > q_j^m$，i 将不会选择 $q_i < R_i(q_j^m) \equiv \underline{q}_i^2$，因 $q_i < \underline{q}_i^2$ 严格劣于 $\underline{q}_i^2, i = 1,2$。故第二轮剔除劣战略后得到 $S_i(2) = \left[R_i(q_j^m), q_i^m\right] \equiv \left[\underline{q}_i^2, q_i^m\right]$，$i \neq j$，$i, j = 1,2$。接着，给定 i 知道 j 不会选择 $q_j < R_j(q_i^m) \equiv \underline{q}_j^2$，$i$ 将不会选择 $q_i > R_i(\underline{q}_j^2) \equiv \underline{q}_i^3$，因 $q_i > \overline{q}_i^3$ 严格劣于 $\overline{q}_i^3$，故第三轮剔除劣战略后得到 $S_i(3) = \left\lfloor \underline{q}_i^2, R_i(\underline{q}_j^2) \right\rfloor \equiv \left\lfloor \underline{q}_i^2, \overline{q}_i^3 \right\rfloor$。如此不断重复剔除下去，每次剔除使剩余的战略空间不断缩小，如图 3.3 所示。一般地，在 $m = 2k+1$ 次剔除时，$S_i(m) = \left\lfloor \underline{q}_i^{2k}, R_i(\underline{q}_j^{2k}) \right\rfloor \equiv \left\lfloor \underline{q}_i^{2k}, \overline{q}_i^{2k+1} \right\rfloor$，$k = 0,1,\cdots$。显然该过程要经过无穷次剔除且序列 $\underline{q}_i^m$ 和 $\overline{q}_i^m$ 都收敛于 q_i^*，$i = 1,2$，$m = 1,\cdots$。故重复剔除的占优战略均衡为 $E(q_1^*, q_2^*)$[7]。

3.2.5 纳什均衡：发现的故事

与重复剔除的占优战略均衡相关的问题还不仅在于前面所提到的对局中人理性程度要求过高及“理性是局中人的共同知识”假定的非现实性问题，一个十分关键的问题是许多博弈模型并不存在重复剔除的占优战略均衡，即许多博弈模型是重复剔除占优不可解的。在博弈论的发展初期，博弈论专家们关注的主要问题是寻找一种均衡概念，这种均衡概念还应保证在非常一般的情况下，博弈模型都存在均衡解。博弈论的创始人之

一冯·诺依曼提出的“最大最小定理”能保证在非常一般的情况下，两人零和博弈总是存在“最大最小均衡”。但是，零和博弈在社会科学中没有多大意义，因为包括经济、军事博弈在内的博弈几乎都是非零和的，即赢方所得并不等于输方所输（两个敌对国家在经过一场核战争后，可能并不存在赢方，结果可能是两败俱伤，即使有胜者，胜者所得也不等于败者所失），甚至可能根本不存在输方。对于超过两人以上的博弈，在冯·诺依曼与莫根斯坦的《博弈论与经济行为》一书中尽管曾作为另一个重要部分，但却没有阐述完全，冯·诺依曼也未能证明对于所有这样的博弈总是存在均衡解。《博弈论与经济行为》的最后 80 页用于研究非零和博弈，但是冯·诺依曼通过引入一个虚构的局中人，用于消费过剩资源或弥补赤字，从形式上将这样的博弈转化为零和博弈。正如一位评论家后来所说：“这个手段确实有帮助，但是并不足以引出非零和博弈的一个完全恰当的处理办法。这是非常遗憾的事情，因为这样的博弈最有可能在实践当中大显身手。”

在纳什提出“纳什均衡”概念之前，冯·诺依曼的零和博弈占据着博弈论研究的中心位置。在 1949 年，兰德（RAND）公司的数学家们、军事战略家们和经济学家们几乎一直在集中研究两人之间完全冲突的两人零和博弈，在这种情况下，我所得即你所失，反之亦然。但是，对于当时苏美两大超级大国间的核冲突问题，零和博弈并不实用。

随着武器变得愈来愈具有破坏性，即便是全面战争也不再形成对手之间毫无共同利益的纯粹冲突的局面。向一个敌人施加最彻底的破坏，用炸弹将他们送回石器时代，将不再是明智之举。正如一些美国战略家在对德作战最后阶段所意识到的那样，当时他们决定不去破坏鲁尔工业区的煤矿和工业设施。

20 世纪 40 年代后期，在作为博弈论的发源地普林斯顿大学，聚集了一大批对博弈论作出重要贡献的数学家和经济学家，他们当时研究博弈论的主要目的是打算在美苏军事冲突中和经济分析中引入严格的数学理论，这种数学理论在描述局中人的战略和行动选择时假定局中人都是理性的，同时，学者们也注意到，还需要引入模型的一个重要变量是局中人之间在决策上的相互依赖性。作为当时另一个博弈论研究中心的兰德公司，也正在考虑如何将博弈论应用于战后的美苏军事平衡研究。兰德的一位核战略家谢林博士曾深刻地指出博弈论需纳入局中人决策互动的必要性，他写道：

在国际事务中，存在相互依存和对抗。两个对手的利益完全相反的纯粹冲突是一个特例，只可能出现在你死我活的完全灭绝性的战争中，在其他类型的战争中也不会出现。相互迁就的可能性与冲突的要素一样重要和富于戏剧性。类似威慑力、有限战争、裁军以及谈判这样的概念，牵涉到可能存在于一场冲突的局中人之间的共同利益和相互依存性。

谢林接着解释了他这样说的原因：“在这些博弈中，尽管冲突的要素提供了值得关注的利益，相互依存性仍然是逻辑结构的一部分，要求某种合作或相互迁就，也许是心照不宣的，也许是明确宣布的，哪怕只是为了阻止共同面临的灾难”（王尔山，2000.10）。

在 1950 年，至少兰德的经济学家们已经认识到，如果博弈论要发展成为一个描述性的理论，可以有效应用于现实生活的军事和经济冲突，人们就必须将注意力集中在同时考虑合作与冲突的博弈。“每个人都已经对零和博弈感到烦燥不安”，经济学家阿罗回忆说，“你正在尝试决定要不要开战，而你又不能说失败者失去的就是胜利者得到的，这真是一件麻烦事。”

对于纳什这样一位正在普林斯顿大学数学系攻读博士学位的满怀雄心壮志的年轻博士生来说，冯·诺依曼理论中的缺陷和瑕疵显得格外诱人，这就像当年一直找不到光波赖以传播的物质以太这一令人困惑的难题吸引过年轻的爱因斯坦一样。纳什立即开始思考由冯·诺依曼和莫根斯坦创立但存在理论和方法问题的博弈均衡问题。在经过 1949 年夏天中一段时间的深思熟虑之后，纳什将几个模糊的想法变得成型和成熟。那年 10 月，他开始感到才思如潮，他终于捕捉到了一种对人类行为的非凡洞察即纳什均衡概念。博士生综合考试结束后几天，纳什就去见冯·诺依曼。先前他曾趾高气扬地对冯·诺依曼的秘书说他想讨论冯·诺依曼教授可能感兴趣的一个问题，对于一个研究生来说，这样做相当大胆莽撞。冯·诺依曼是一位公众人物，在偶尔进行的演讲之外与普林斯顿的研究生没有什么接触，通常也不鼓励他们前来请教研究中遇到的问题。

冯·诺依曼端坐在一张巨大的桌子旁边，穿着昂贵的三件套西装，打了丝质领带，口袋里露出精致的手帕，整个人看上去与其说像个学者，倒不如说更像一个富有的银行总裁。他也确实和公务繁忙的行政人员一样心事重重，当时正担任着几个顾问职务，“没完没了地和奥本海默争论”氢弹研制的问题，同时还指导两台电子计算机样机的建造和程序编制工作。他做了一个手势让纳什坐下。他知道纳什是谁，只是对于他的来访感到有点儿困惑。

他细心地听着，微微抬头，手指轻轻敲着桌面。纳什开始描述自己已经想好的证明当局中人多于两人的情况下博弈均衡存在的方法以及相应的纳什均衡概念。不过，没等他说完几个互不相关的句子，冯·诺依曼突然打断了他的话，在纳什尚未说到这个证明的结论之前抢先说：“小菜一碟，你知道，这只不过是一个不动点定理。”

纳什观察到，在博弈互动中，局中人选择的战略给局中人带来的支付不仅取决于自己所选择的战略，而且还同时取决于其他局中人所选择的战略，同时，局中人之间在博弈中不一定是竞争性的关系，更多的还存在着合作性关系，即局中人之间可能在某些战略组合上达成一种互利性的共赢或合作。这样，局中人选择什么样的战略对于局中人是最优的取决于其他局中人的战略选择，反之亦然。纳什由此想到，如果给定一个局中人的战略组合，其中任一局中人选定的战略在给定其他局中人战略选择下是最优战略，那么，可预测此时没有任何一位局中人有改变已选定战略的动机。纳什将这种“胶着”状态定义为一种均衡，即纳什均衡。纳什还进一步证明，在非常一般的情况下，博弈模型总是存在至少一个纳什均衡。

在与冯·诺依曼的会见后过了几天，纳什向他的朋友盖尔（Gale）谈到了他的发现：“我觉得我已经找到了一个办法，可以将冯·诺依曼的最大最小定理加以普遍化。……整个理论就是建立在这个基础上。无论局中人数多少都适用，也不仅仅限于零和博弈。……我会将这称为一个均衡点”。均衡的概念就是，这是一个自然的静止点，倾向于持续下去。盖尔比冯·诺依曼更加深刻地看出了纳什观点的意义。“唔”，他说，“这是一个不错的命题。”盖尔意识到纳什的想法能够应用于更加广泛的真实世界环境中，远远超过了冯·诺依曼的零和架构。“他有一个想法可以运用到裁军方面。”盖尔后来说。不过，最让盖尔入迷的还不是纳什的主意可能有些什么用处，却在于它的优美和普遍性。他说：“其中的数学非常优美，在数学上简直是太正确了”。

盖尔积极促动纳什将结果正式发表。盖尔回忆说：“我说这是一个了不起的结果，应该优先处理。”他告诉纳什，他可以肯定纳什已经有了一个绝妙的命题，同时也敦促纳什抢在别人想到一个相似的主意之前尽快将这个成果纳入自己的名下。盖尔建议请美国科学院的一名院士帮忙，将这个证明递交科学院每月出版的学报。“他古怪莫测，也许永远不会想到要这样做。”盖尔曾说，“因此他把他的证明交给我，我起草了写给国家科学院的备忘录。”普林斯顿大学数学系主任莱夫谢茨（Lefschety）立即将这份备忘录递交科学院，结果在 1950 年 11 月的学院学报上刊登出来。盖尔后来补充说：“我当然马上就看出这就是一篇博士学位论文，但是我那时并不知道它会是一个诺贝尔奖得奖作品。”

博弈论的整个大厦建立在两个定理之上，它们分别是冯·诺依曼 1928 年提出的最大最小定理和纳什 1950 年发表的均衡定理。纳什定理可以看做冯·诺依曼定理的一种推广，就像纳什所做的那样，不过，它同时还是一种彻底的决裂。冯·诺依曼定理是他有关纯粹对立的所谓两人零和博弈的理论的奠基石。但是，两人零和博弈与现实世界实际上没有多少联系，即便是在战争期间，也几乎总是可以通过合作得到某些好处。纳什引入了合作博弈与非合作博弈的区别。合作博弈当中，局中人可以与其他局中人订立可执行的协议。换句话说，作为一个团体，他们可以完全将自己托付给一些具体的战略。与此相反，在一个非合作博弈当中，这样的集体委托绝对不可能出现，这里没有可执行的协议。通过将这个理论扩展到牵涉各种合作和竞争的博弈，纳什成功地打开了将博弈论应用到经济学、政治学、社会学乃至进化生物学的大门。

虽然纳什采取的战略形式与冯·诺依曼提出的想法相同，他的解决问题的方式却完全不一样。在冯·诺依曼和莫根斯坦的那部著作里，超过一半篇幅用于论述合作理论。另外，冯·诺依曼和莫根斯坦的解的概念，也就是那个被叫做独立解的东西[8]，其实并不存在于每一个博弈里。与此相反，纳什在自己论文的第 6 页证明，每个非合作博弈，只要局中人的数目和他们可选择的纯战略的数目都有限，就都有至少一个（混合战略）纳什均衡。[9]

如今，纳什的源自战略博弈的均衡概念是社会科学和生物学相关理论的基本范式之一。纳什的洞察力在很大程度上使博弈论像《新帕尔格雷夫经济学大辞典》一书中所说的那样，是“处理一个日益复杂化学科的有力而优美的方法，好比天体力学的牛顿方法取代了古人那些原始且专门的方法一样”。如同牛顿的引力理论和达尔文的自然选择理论等许多伟大科学思想一样，纳什的想法起初也显得太简单而没有意思、太狭隘而没有广泛用途，后来又变成太显而易见，以至于人们相信它不可避免总会被某人发现。正如与纳什和海萨尼分享 1994 年度诺贝尔经济学奖的德国经济学家泽尔腾所说：“从总体来看，没有人能够预见到纳什均衡会给经济学和社会科学带来这样巨大的影响，更不必说指望纳什的均衡概念会对生物学的理论具有重要意义”。人们没有立即意识到它的重要性，就连它的 21 岁作者本人也不例外，当然还有那位激励了纳什的天才人物——冯·诺依曼。

纳什均衡所揭示出的深刻经济学道理也唤醒了沉迷于新古典经济学的经济学家们，因为各个局中人根据自己的最优战略行事，同时估计其他局中人也按照他们的最优战略行事，从这一组局中人的观点来看，不一定就是最优的解决方案。譬如，“囚徒困境”就与经济学中斯密的“看不见的手”的隐喻互相矛盾。处于博弈中的各个局中人都在追求自己的个人利益，但他们不一定会增进整个集体的利益。现代经济学由此开始注意到个人理性和集体理性的矛盾与冲突，但是解决这个问题的办法并不是像传统经济学主张的那样通过政府干预来避免市场失灵所导致的无效状态，而是认为，如果一种制度安排不能满足个人理性的话，就不可能实行下去。所以解决个人理性与集体理性之间冲突的办法不是否认个人理性，而是设计一种机制，在满足个人理性的前提下达到集体理性。认识到个人理性与集体理性的冲突对于认识制度安排是非常重要的。[10]

3.2.6 纳什均衡：定义与概念

定义 3.6　对于 n 人战略式表述博弈 $G=\{S_1,\cdots,S_n;u_1,\cdots,u_n\}$，若战略组合 $s^*=(s_1^*,\cdots,s_n^*)$ 满足如下条件，则称 s^* 是一个纳什均衡：

$u_i(s_i^*,s_{-i}^*)\geq u_i(s_i,s_{-i}^*)$，$\forall s_i\in S_i, i=1,\cdots,n$（符号“$\forall$”表示“任意的”）

即是说如果对于每一个 $i=1,\cdots,n$，s_i^* 是给定其他局中人选择 $s_{-i}^*=(s_1^*,\cdots,s_{i-1}^*,s_{i+1}^*,\cdots,s_n^*)$ 的情况下第 i 个局中人的最优战略。

或者用另一种表达方式：当且仅当 s_i^* 是下述最大化问题的解时，s^* 是一个纳什均衡：

$$s_i^*\in\arg\max_{s_i\in S_i} u_i(s_1^*,\cdots,s_{i-1}^*,s_i,s_{i+1}^*,\cdots,s_n^*)，\quad i=1,\cdots,n$$

纳什均衡的含义是说：当局中人在某一选定的战略组合下都没有积极性偏离各自已选定的战略时，该战略组合就构成一个纳什均衡。

定义 3.6 给出的纳什均衡是一种“弱纳什均衡”概念。当定义 3.6 中的不等式为严

格不等式时，我们得到“强纳什均衡概念”，于是我们有如下定义：

定义 3.7　一个纳什均衡是强的（Strict 或 Strong），如果给定其他局中人的战略，每一个局中人的最优选择是唯一的。即是说s^*是一个强纳什均衡，当且仅当对于所有的$i=1,\cdots,n$，$s_i^{'} \neq s_i^*$，有

$$u_i(s_i^*, s_{-i}^*) > u_i(s_i^{'}, s_{-i}^*)$$

如果一个纳什均衡是强的，则没有任何局中人在均衡战略与某些其他战略之间是无差异的，在弱纳什均衡情况下，有些局中人可能在均衡战略与非均衡战略之间是无差异的，既然是无差异的，为什么选择均衡战略而不选择其他战略呢？由于这个原因，强纳什均衡比弱纳什均衡是一个更为可取的概念。[11]此外，强纳什均衡对博弈支付的微小变化并不敏感，这保证了博弈模型预测功能的稳健性。

3.2.7 用划线法寻找战略式表述纯战略纳什均衡

纳什存在性定理保证有限博弈都至少存在一个纳什均衡，但对于一个具体的博弈模型，我们可能更感兴趣的是如何将纳什均衡求解出来。一般而论，不同情况下的纳什均衡有不同的求解方法。当战略空间是连续性的且支付函数充分充滑时，通常运用微分学方法求解纳什均衡，但在两人有限博弈的战略式表述场合，我们可直接使用一种十分简便的方法找出模型中所有的纯战略纳什均衡，这就是“划线法”。在表 3.12 中，我们在给定 A 的每一个战略选择下找到 B 的最大支付所对应的 B 的战略，然后在该最大支付的下端划上一条短横线；同样地，我们接着又在给定 B 的每一战略选择下找到 A 的最大支付所对应的 A 的战略，然后在该最大支付的下端划上一条短横线。最后，我们将那些 A 和 B 的支付下端都划有短横线所对应的战略组合找出来，它们就是纯战略纳什均衡；于是，战略组合（D, M）构成了一个纳什均衡。

表 3.12　用划线法求解纳什均衡

		B: L	B: M	B: R
A	U	3, 2	4, <u>7</u>	5, 1
A	H	<u>6</u>, 1	2, <u>8</u>	1, 1
A	D	3, 7	<u>8</u>, <u>9</u>	<u>10</u>, 4

根据纳什均衡的定义，我们很容易理解为什么通过划线法找出的战略组合是纳什均衡。在两个局中人的支付下端都划有短横线的战略组合中，给定任一局中人的战略，另一个局中人的战略都是最优战略，因而这个战略组合就是一个纳什均衡。

3.2.8 混合战略博弈纳什均衡

请读者回忆你的孩提时代，想想你们那时曾玩过哪些游戏，我们想少不了有一种游戏大家都曾玩过，那就是“石头、剪子、布”。这种游戏相信有着古老的起源，甚至世界各地都有类似玩法。自上个世纪末以来，一种被称为“棍子、老虎、鸡、虫子”的新玩法在较成年的年轻人中盛行开来，这种游戏与“石头、剪子、布”游戏大同小异，但主要在一些酒宴上作为取代变化太多且略显粗鲁的旧式中国猜拳令而玩，据说这种游戏特别受到宴会上女士们的欢迎。在“石头、剪子、布”游戏中，有两个玩手，游戏规则是这样的：石头胜剪子，剪子胜布，布胜石头。譬如，当一方出石头而另一方出布时，则后者就赢了前者。在“棍子、老虎、鸡、虫子”游戏中，也有两个玩手，游戏规则是：棍子胜（打死）老虎，老虎胜（吃掉）鸡，鸡胜（吃掉）虫子，而虫子胜（蛀掉）棍子。我们现在应该想到，这些著名的游戏实际上是典型的博弈。“石头、剪子、布”和“棍子、老虎、鸡、虫子”都是二人零和博弈。对于二人零和博弈，冯・诺依曼提出的最大最小定理表明它们通常有最大最小均衡，这是否说博弈中的局中人真的会选择最大最小均衡战略吗？让我们一起来回忆儿时玩“石头、剪子、布”时的情形，显然，我们（至少绝大多数人）在儿时玩这类游戏时根本未想到过选择最大最小均衡战略，恐怕连最大最小均衡概念本身即使在大多数（未学过博弈论的）成年人中也未必被大家所知道。我们在表 3.13 中给出“石头、剪子、布”的一种战略式表述。

表 3.13 “石头、剪子、布”博弈

		B 石头	剪子	布
	石头	0, 0	1, –1	–1, 1
A	剪子	–1, 1	0, 0	1, –1
	布	1, –1	–1, 1	0, 0

显然，该博弈有九个最大最小均衡，即局中人 A 和 B 的所有战略都是最大最小均衡战略，因此，我们若按最大最小均衡战略来决定局中人的战略选择，则局中人选任意战略都是无差异的。但是，当局中人 A 选择“石头”时，B 会选择“布”，而给定 B 选“布”，A 的最优选择又是“剪子”而不是“石头”，而给定 A 选“剪子”，B 的最优选择又是“石头”而非“布”，……。在“石头、剪子、布”游戏中，不存在纯战略纳什均衡，但最大最小均衡既与我们的实际经验相左，又存在上述的不确定性问题。那么，这种博弈是否存在均衡呢？相信大家在玩“石头、剪子、布”游戏时都有这样一种体验——为了避免对方猜出自己会选择什么，我们总是在出示自己的选择之前尽量迷惑对手，不给予对方哪怕是一丁点会暴露自己选择意向的信息。在重复多次的游戏玩演过程

中，我们尽量不重复或少重复自己的同一战略选择。一言以概之，我们在玩“石头、剪子、布”时，总是尽量随机性地选择战略——在“石头、剪子、布”中随机地选择。那么，这是一种什么博弈，它是否也存在相应的均衡概念——这将是我们在本小节中将要介绍的混合战略博弈（mixed strategy game）概念，其相应的均衡概念是纯战略纳什均衡概念在混合战略博弈中的扩充，即混合战略博弈纳什均衡（MNE）。

3.2.9 混合战略博弈：概念与定义

正如我们在前一小节中所看到的那样，“石头、剪子、布”游戏博弈没有纯战略纳什均衡，除了这个例子之外，我们还可以轻易地举出许多不存在纯战略纳什均衡的例子，见下面给出的几个例子。

例 3.11 “棍子·老虎·鸡·虫子”博弈

这个博弈在前一小节中已作了介绍，这里用具体的博弈矩阵来加以刻画，见表 3.14。

显然，这个博弈没有纯战略纳什均衡（它也是一个零和博弈）。

表 3.14 “棍子·老虎·鸡·虫子”博弈

		2			
		棍子	老虎	鸡	虫子
1	棍子	0, 0	1, –1	0, 0	–1, 1
	老虎	–1, 1	0, 0	1, –1	0, 0
	鸡	0, 0	–1, 1	0, 0	1, –1
	虫子	1, –1	0, 0	–1, 1	0, 0

例 3.12 社会福利博弈

政府在实行公共福利政策时，通常会遇到激励问题，譬如，当政府对失业者进行救济时，可能会激励失业者不再愿意努力寻找工作。表 3.15 给出了这种博弈的一种具体描述。

表 3.15 社会福利博弈

		失业者	
		寻找工作	游荡
政府	救济	3, 2	–1, 3
	不救济	–1, 1	0, 0

政府救济失业者的前提是失业者必须试图寻找工作，否则视失业者为自愿失业者，对于自愿失业者政府不会给予救济；失业者只有在得不到政府救济时才会寻找工作，否

则会饿饭的。这个博弈不存在纯战略纳什均衡。给定政府救济，失业者的最优战略是游手好闲；给定流浪汉无所事事地四处游荡，政府的最优战略是不救济；给定政府不救济，失业者的最优战略是寻找工作；而给定失业者寻找工作，政府的最优战略是救济；……如此等等，没有一个战略组合构成纯战略纳什均衡。

例 3.13　猜谜游戏（matching pennies）

两个小孩子手里各拿着一枚硬币，决定要出示正面或反面。如果两枚硬币同时以相同的面被两个小孩子出示，则小孩子 A 付给小孩子 B 一分钱；否则，小孩子 B 输给小孩 A 一分钱。表 3.16 给出了这个博弈的博弈矩阵。

表 3.16　猜谜游戏

		小孩 B 正面	小孩 B 反面
小孩 A	正面	–1, 1	1, –1
	反面	1, –1	–1, 1

这也是一个没有纯战略纳什均衡的纯战略博弈。

象这类不存在纯战略纳什均衡的博弈在现实生活中是广泛存在的，诸如扑克牌比赛、橄榄球赛和战争等都是常见的例子。那么，在这种没有纯战略纳什均衡的博弈中，局中人是按什么样的策略进行战略选择的呢？我们观察到，譬如，在“石头、剪子、布”博弈中，局中人总是在选定其战略前试图猜中对手的战略选择（尽管常常是徒劳的），同时，局中人又总是力图避免自己的战略被对方猜中。避免自己的战略被对方猜中的策略是尽量随机性地选择自己的战略，但这种随机性并不意味着是一种均匀分布或接近均匀分布的战略选择概率，因为自己的战略选择不仅要避免自己的战略选择被对方猜透，而且还要根据自己对对方战略选择的事前预测及根据不同战略组合带来的不同支付情况来作出最优的战略选择。由此，我们得到所谓“混合战略博弈”（mixed strategy game）的概念。

定义 3.8　n 人战略式表述博弈 $G=\{S_1,\cdots,S_n;u_1,\cdots,u_n\}$ 中，$S_i=\{s_{i1},\cdots,s_{ik(i)}\}$，概率密度 $\sigma_i=(\sigma_{i1},\cdots,\sigma_{ik(i)})$ 被称为局中人 i 的一个混合战略，$\sigma_{ik}=\sigma_i(s_{ik})$ 是 i 选择 s_{ik} 的概率，$0\le\sigma_{ik}\le 1$，$\sum_{k=1}^{k(i)}\sigma_{ik}=1$，$k=1,\cdots,k(i),i=1,\cdots,n$，$k(i)$ 为 S_i 中纯战略个数（$k(i)$ 可为无穷大）。

显然，混合战略是纯战略概念的一种扩充，因为当某个 $\sigma_{ik}=1(\sigma_{ik'}=0,k\neq k',k,k'=1,\cdots,k(i))$ 时，混合战略就“退化”为一种纯战略。当局中人按混合战略进行博弈时，我们称博弈为混合战略博弈。

记局中人i可选择的所有混合战略（当然包括所有的纯战略）构成的集合为$\sum_i$，称$\sum_i$为i的混合战略空间，$\sigma_i \in \sum_i$；$\sigma = (\sigma_1, \cdots, \sigma_n)$为混合战略组合（mixed strategy profile）；笛卡儿积$\sum = \prod_{i=1}^{n} \sum_i$为混合战略组合空间，$\sigma \in \sum$。读者可能会注意到，上述定义是不完全的，因为它未说明当局中人按混合战略进行博弈时局中人所获得的支付是什么。我们在下文中将立即来解决这一问题。

在混合战略博弈中，任一局中人不能准确判定博弈的最终纯战略组合是什么，因为至少其他局中人的纯战略选择对他来说只是一种不太准确的预测，所以，局中人对其所获得的支付的判断只能是一种预期值或期望效用。记$v_i(\sigma) = v_i(\sigma_i, \sigma_{-i})$（这里$-i$表示除$i$之外的其他局中人，即$\sigma_{-i} = (\sigma_1, \cdots, \sigma_{i-1}, \sigma_{i+1}, \cdots, \sigma_n)$）为局中人$i$的期望支付函数。于是，我们有

$$v_i(\sigma_i, \sigma_{-i}) = \sum_{s \in S} p(s) u_i(s) \qquad (3.6)$$

其中$p(s)$为s作为博弈结果的战略组合的概率。在完全信息静态博弈中，我们有理由假定各个局中人在选择纯战略时是相互独立地作出决策的，因而就有$p(s) = \prod_{j=1}^{n} \sigma_j(s_j)$，于是得到

$$v_i(\sigma_i, \sigma_{-i}) = \sum_{s \in S} \prod_{j=1}^{n} \sigma_j(s_j) u_i(s) \qquad (3.7)$$

3.2.10 混合战略博弈纳什均衡：定义与解法

下面，我们定义混合战略博弈纳什均衡如下：

定义 3.9　n人战略式表述博弈$G = \{S_1, \cdots, S_n; u_1, \cdots, u_n\}$中，混合战略组合$\sigma^* = (\sigma_1^*, \cdots, \sigma_i^*, \cdots, \sigma_n^*)$是一个纳什均衡，如果对于所有的$i = 1, \cdots, n$，有：

$$v_i(\sigma_i^*, \sigma_{-i}^*) \geq v_i(\sigma_i, \sigma_{-i}^*), \forall \sigma_i \in \sum_i \qquad (3.8)$$

混合战略博弈作为纯战略博弈的自然扩充是必要的，因为在这种扩充下才有著名的纳什存在性定理（我们在第四章中介绍纳什存在性定理），也就是说，纳什存在性定理保证在非常一般的情况下存在的纳什均衡包括仅存在混合战略博弈纳什均衡的情形。

关于混合战略博弈纳什均衡的求解，我们有如下定理及推论。

定理 3.2　若$\sigma_i = (\sigma_{i1}, \cdots, \sigma_{ik(i)})$是相对于给定$\sigma_{-i}$的一个最优混合战略，则对于所有的$\sigma_{ik} > 0$，有

$$v_i(s_{ik}, \sigma_{-i}) \geq v_i(s_{ik'}, \sigma_{-i}), \forall s_{ik'} \in S_i, \quad i = 1, \cdots, n$$

证明：这是因为，若不然，存在某$s_{ik'}$使$v_i(s_{ik'}, \sigma_{-i}) > v_i(s_{ik}, \sigma_{-i})$，此时用$s_{ik'}$取代原最优混合战略中的$s_{ik}$后局中人$i$的支付会提高，但这与原混合战略的最优性相矛盾，故是不可能的。

事实上，若$\sigma_i = (\sigma_{i1}, \cdots, \sigma_{ik}, \cdots, \sigma_{ik(i)})$是最优混合战略，则

$v_i(\sigma_i,\sigma_{-i}) \ge v_i(\sigma_i^{'},\sigma_{-i}), \forall \sigma_i^{'} \in \sum_i$， $i=1,\cdots,n$

即

$$\sum_{s\in S}\prod_j \sigma_j(s_j)u_i(s) \ge \sum_{s\in S}\prod_j \sigma_j^{'}(s_j)u_i(s)$$

$$\sum_{s\in S}\sigma_i(s_i)\prod_{j\ne i}\sigma_j(s_j)u_i(s) \ge \sum_{s\in S}\sigma_i^{'}(s_i)\prod_{j\ne i}\sigma_j^{'}(s_j)u_i(s)$$

上式左端为 $\sum_{s_{ik}\in S_i}\sigma_i(s_{ik})\sum_{\substack{s\in S\\ s_i=s_{ik}}}\prod_{j\ne i}\sigma_j(s_j)u_i(s)=\sum_k \sigma_{ik}v_i(s_{ik},\sigma_{-i})$

同理，上式右端为 $\sum_k \sigma_{ik}^{'}v_i(s_{ik},\sigma_{-i})$

令 $\sigma_{ik'}^{'}=\sigma_{ik}+\sigma_{ik'}, \sigma_{ik}^{'}=0$

即 $\sigma_{i1},\cdots,\sigma_{ik'},\cdots,\sigma_{ik},\cdots,\sigma_{ik(1)},\rightarrow \sigma_{i1}^{'},\cdots,\sigma_{ik(i)}^{'}$

+

$=\sigma_{i1},\cdots,\sigma_{ik'}+\sigma_{ik},\cdots,\sigma_{ik-1},0,\sigma_{ik+1},\cdots,\sigma_{ik(i)}$

则以上不等式为

$\sigma_{ik}v_i(s_{ik},\sigma_{-i})+\sigma_{ik'}v_i(s_{ik'},\sigma_{-i})+\sum_{l\ne k',k}\sigma_{il}v_i(s_{il},\sigma_{-i})$

$\ge(\sigma_{ik}+\sigma_{ik'})v_i(s_{ik'},\sigma_{-i})+\sum_{l\ne k',k}\sigma_{il}v_i(s_{il},\sigma_{-i})$

即 $\sigma_{ik}v_i(s_{ik},\sigma_{-i})+\sigma_{ik'}v_i(s_{ik'},\sigma_{-i})\ge(\sigma_{ik}+\sigma_{ik'})v_i(s_{ik'},\sigma_{-i})$

得到 $v_i(s_{ik},\sigma_{-i})\ge v_i(s_{ik'},\sigma_{-i})$

这说明所有以正的概率进入最优混合战略的纯战略都是最优战略。因为期望支付是一系列纯战略期望支付的凸组合，将不等式右端换成混合战略期望支付时不等式仍成立。

证毕。

推论 3.2.1　在定理 3.2 中，若 $\sigma_{il}>0$， $\sigma_{il'}>0$，则必有 $v_i(s_{il},\sigma_{-i})=v_i(s_{il'},\sigma_{-i})$ $=v_i(\sigma_i,\sigma_{-i})$，即所有以正的概率进入最优混合战略的纯战略之间必定是无差异的。

证明：否则，存在 $m,\sigma_{im}>0$，使

$v_i(s_{im},\sigma_{-i})<\max_{l|\sigma_{il}>0}v_i(s_{il},\sigma_{-i})\overset{\Delta}{=}v_i(s_{iM},\sigma_{-i})$

令 $\sigma_{im}^{'}=0$， $\sigma_{iM}^{'}=\sigma_{im}+\sigma_{iM}$， $\sigma_{il}^{'}=\sigma_{il}$， $l\ne m,M$，

则 $v_i(\sigma_i^{'},\sigma_{-i})=\sum\sigma_{ik}^{'}v_i(s_{ik},\sigma_{-i})>v_i(\sigma_i,\sigma_{-i})$，这与 σ_i 的最优性矛盾。

当证明了第一个等式成立后，最后一个等式是显然的。

证毕！

这说明：在最优混合战略 σ_i 中以正概率选择的纯战略 s_{ik} 必然是最优纯战略，同时又必然是最优混合战略，因

$v_i(s_{ik},\sigma_{-i})\ge v_i(s_{ik'},\sigma_{-i}), \forall s_{ik'}\in S_i$

故$\sigma_{ik'}v_i(s_{ik},\sigma_{-i}) \ge \sigma_{ik'}v_i(s_{ik'},\sigma_{-i}), 0 \le \sigma_{ik'} \le 1$，在不等式两边求和：

$(\sum\sigma_{ik'})v_i(s_{ik},\sigma_{-i}) \ge \sum\sigma_{ik'}v_i(s_{ik'},\sigma_{-i}) = v_i(\sigma_i^{'},\sigma_{-i}), \forall\sigma_i^{'} \in \sum_i$

因$\sum\sigma_{ik'} = 1$

故$v_i(s_{ik},\sigma_{-i}) \ge v_i(\sigma_i^{'},\sigma_{-i}), \forall\sigma_i^{'} \in \sum_i$

推论 3.2.2　在定理 3.2 中，若局中人有k个纯战略是最优的，则任何以正的概率选择其中一些或所有这些纯战略的混合战略也是最优的。

证明：不妨设$s_{i1},s_{i2},\cdots,s_{ik}$是最优纯战略，用上述类似方法可证它们的概率组合也是最优混合战略，即

$v_i(s_{il},\sigma_{-i}) \ge v_i(\sigma_i^{'},\sigma_{-i}), \forall\sigma_i^{'} \in \sum_i, l = 1,\cdots,k$

令$\sum_l \sigma_{il} = 1$，$0 \le \sigma_{il} \le 1$，$l = 1,\cdots,k$，$\sigma_{ih} = 0, h = k+1,\cdots,k(i)$

则$\sum_{h=1}^{k(i)}\sigma_{ih}v_i(s_{ih},\sigma_{-i}) \ge \left(\sum_l \sigma_{il}\right)v_i(\sigma_i^{'},\sigma_{-i}) = v_i(\sigma_i^{'},\sigma_{-i}), \forall\sigma_i^{'} \in \sum_i$

证毕！

推论 3.2.3 若$\sigma^* = (\sigma_1^*,\cdots,\sigma_n^*)$是混合战略纳什均衡，则必有

$v_i(\sigma_i^*,\sigma_{-i}^*) \ge v_i(s_i,\sigma_{-i}^*), \forall s_i \in S_i$，

反之，当$v_i(\sigma_i^*,\sigma_{-i}^*) \ge v_i(s_i,\sigma_{-i}^*)$，对任何$s_i \in S_i$成立时，有

$v_i(\sigma_i^*,\sigma_{-i}^*) \ge v_i(s_{il},\sigma_{-i}^*)$，$l = 1,\cdots,k(i)$

于是有$v_i(\sigma_i^*,\sigma_{-i}^*) \ge v_i(\sigma_i,\sigma_{-i}^*)$，$\forall\sigma_i \in \sum_i$。

证明：只需证第一个结论成立即可。对任何$0 \le \sigma_{il} \le 1$，$\sum\sigma_{il} = 1$，若已有$v_i(\sigma_i^*,\sigma_{-i}^*) \ge v_i(s_i,\sigma_{-i}^*), \forall s_i \in S_i$，则有

$\sigma_{il}v_i(\sigma_i^*,\sigma_{-i}^*) \ge \sigma_{il}v_i(s_{il},\sigma_{-i}^*)$

两边求和：$v_i(\sigma_i^*,\sigma_{-i}^*) \ge \sum\sigma_{il}v_i(s_{il},\sigma_{-i}^*) = v_i(\sigma_i,\sigma_{-i}^*), \forall\sigma_i \in \sum_i$，即$\sigma_i^*$是混合战略纳什均衡；推论中第一个结论是显然的。

证毕！

因此，混合战略纳什均衡σ^*也可定义为：

定义 3.10 若σ^*满足：

$$v_i(\sigma_i^*,\sigma_{-i}^*) \ge v_i(s_i,\sigma_{-i}^*), \forall s_i \in S_i \tag{3.9}$$

则σ^*是一个混合战略纳什均衡。

下面，我们来看看如何求解一个给定混合战略博弈的纳什均衡。基于推论 3.2.1，我们可以求解某些（主要是$k(i) = 2$的二人博弈）混合战略博弈的纳什均衡，下面用例子说明。

例 3.14　社会福利博弈的纳什均衡求解

我们回看例 3.12 中给出的社会福利博弈，根据表 3.14，我们易知该博弈不存在纯战略纳什均衡（只需运用划线法就可证实这一点）。我们下面来看看该博弈有没有混合战

略博弈纳什均衡。根据我们在下一章中即将介绍的纳什存在性定理（在前面已多次提及），该博弈至少存在一个混合战略博弈纳什均衡，我们现在的任务是将这样的均衡找出来。设政府的混合战略为$\sigma_1=(\theta,1-\theta)$，其中$\theta$为政府对失业者进行救济的概率；失业者的混合战略为$\sigma_2=(\gamma,1-\gamma)$，其中$\gamma$为失业者找工作的概率。

根据推论 3.2.1 及表 3.14 有

$v_i(\text{救济},\sigma_2)=v_i(\text{不救济},\sigma_2)$

即$3\gamma+(-1)(1-\gamma)=(-1)\gamma+0\cdot(1-\gamma)$

得$\gamma=\dfrac{1}{5}=0.2$

同样有

$v_2(\sigma_1,\text{寻找工作})=v_2(\sigma_1,\text{游荡})$

即$2\theta+1\cdot(1-\theta)=3\theta+0\cdot(1-\theta)$

得$\theta=\dfrac{1}{2}=0.5$

故混合战略博弈纳什均衡为$\{(0.5,0.5),(0.2,0.8)\}$

例 3.15　猜谜游戏的纳什均衡求解

前述例 3.13 中给出的猜谜游戏的混合战略博弈纳什均衡是什么呢？当然，我们完全可以运用类似于前例（例 3.14）中的方法将其求解出来，但实际上还有比其更为简便或直接的方式获得纳什均衡，因为从表 3.15 可以看出，博弈中的两个局中人之间完全是对称的，并且局中人的所有战略之间也是对称的，因而纳什均衡解也可能是对称的，所以我们可以直接猜出其均衡为

$\{(0.5,0.5),(0.5,0.5)\}$

读者不妨运用类似于例 3.14 中的方法进行计算，可证实这种猜测是正确的，并可证明除此均衡外别无其他均衡。

例 3.16 “石头、剪子、布”博弈的纳什均衡求解

显然，类似于例 3.15，我们发现这个博弈也是对称的（见表 3.15），因而猜测其混合战略博弈纳什均衡必然包括一种对称型的均衡，即$\left\{\left(\frac{1}{3},\frac{1}{3},\frac{1}{3}\right),\left(\frac{1}{3},\frac{1}{3},\frac{1}{3}\right)\right\}$。除此均衡之外，是否还有别的均衡？下面，我们通过推证来证明该博弈只有这一个均衡。

设$\sigma_A=(\theta_1,\theta_2,1-\theta_1-\theta_2),\sigma_B=(\gamma_1,\gamma_2,1-\gamma_1-\gamma_2)$

其中$\theta_1\geq0$，$\theta_2\geq0$，$\gamma_1\geq0$，$\gamma_2\geq0$，$\theta_1+\theta_2\leq1$，$\gamma_1+\gamma_2\leq1$，据式（3.7），我们有

$$\begin{aligned}v_A&=\theta_1[0\cdot\gamma_1+1\cdot\gamma_2+(-1)(1-\gamma_1-\gamma_2)]+\theta_2[(-1)\cdot\gamma_1+0\cdot\gamma_2+1\cdot(1-\gamma_1-\gamma_2)]\\&\quad+(1-\theta_1-\theta_2)[1\cdot\gamma_1+(-1)\cdot\gamma_2+0\cdot(1-\gamma_1-\gamma_2)]\\&=\theta_1(\gamma_1+2\gamma_2-1)+\theta_2(-2\gamma_1-\gamma_2+1)+(1-\theta_1-\theta_2)(\gamma_1-\gamma_2)\end{aligned}$$

记 $\gamma = \max[(\gamma_1 + 2\gamma_2 - 1), (-2\gamma_1 - \gamma_2 + 1), (\gamma_1 - \gamma_2)]$

若$(\gamma_1 + 2\gamma_2 - 1), (-2\gamma_1 - \gamma_2 + 1), (\gamma_1 - \gamma_2)$三个数中只有某一个数等于$\gamma$，则给定$\gamma_1, \gamma_2$下$A$的最优战略是纯战略，而我们已知该博弈没有纯战略纳什均衡，故必须假定这三个数中至少有两个数等于γ，这有四种情形：

a. $\gamma_1 + 2\gamma_2 - 1 = -2\gamma_1 - \gamma_2 + 1 = \gamma > \gamma_1 - \gamma_2$

b. $-2\gamma_1 - \gamma_2 + 1 = \gamma_1 - \gamma_2 = \gamma > \gamma_1 + 2\gamma_2 - 1$

c. $\gamma_1 + 2\gamma_2 - 1 = \gamma_1 - \gamma_2 = \gamma > -2\gamma_1 - \gamma_2 + 1$

d. $\gamma_1 + 2\gamma_2 - 1 = -2\gamma_1 - \gamma_2 + 1 = \gamma_1 - \gamma_2 = \gamma$

在情形 a. $\gamma_1 < \frac{2}{3}$，$\gamma_2 = \frac{2}{3} - \gamma_1$

在情形 b. $\gamma_1 = \frac{1}{3}$，$\gamma_2 < \frac{1}{3}$

在情形 c. $\gamma_1 > \frac{1}{3}$，$\gamma_2 = \frac{2}{3}$

在情形 d. $\gamma_1 = \frac{1}{3}$，$\gamma_2 = \frac{1}{3}$

下面分析在给定γ_1, γ_2在上述不同数值范围内取值时局中人 A 的最优θ_1, θ_2取值。

在情形 a，应有$1 - \theta_1 - \theta_2 = 0$，即$\theta_1 + \theta_2 = 1$，且$\theta_1 \cdot \theta_2 \neq 0$（否则局中人 A 的战略就退化成了纯战略），于是 A 的最优θ_1, θ_2取值在如下范围：

$\theta_2 = 1 - \theta_1$， $\theta_1 \in (0,1)$

在情形 b，应有θ_1=0，$\theta_2 + (1 - \theta_1 - \theta_2) = 1$且$\theta_2(1 - \theta_1 - \theta_2) \neq 0$。于是 A 的最优$\theta_1, \theta_2$取值在如下范围：

$\theta_1 = 0$， $\theta_2 \in (0,1)$

在情形 c，应有$\theta_2 = 0$，$\theta_1 + (1 - \theta_1 - \theta_2) = 1$且$\theta_1(1 - \theta_1 - \theta_2) \neq 0$，于是 A 的最优$\theta_1, \theta_2$取值在如下范围：

$\theta_2 = 0$， $\theta_1 \in (0,1)$

在情形 d，应有$\theta_1 \neq 1$，$\theta_2 \neq 1$，$1 - \theta_1 - \theta_2 \neq 1$，即

$\theta_1 \neq 1$，$\theta_2 \neq 2$，$\theta_1 + \theta_2 > 0$

现在，我们来看当局中人 A 分别选取上述不同范围内的θ_1, θ_2时，局中人 B 的最优选择γ_1, γ_2是否还是原来决定最优θ_1, θ_2的那些γ_1, γ_2，如果是这样的，则可得到一些均衡，否则就不能构成均衡。注意，由对称性，给定θ_1, θ_2时γ_1, γ_2的最优选择分析与给定γ_1, γ_2时θ_1, θ_2的最优选择分析是类似的。

当θ_1, θ_2满足：

$\theta_2 = 1 - \theta_1$，$\theta_1 \in (0,1)$

仅当$\theta_1 = \frac{2}{3}$时，θ_1, θ_2取类似于上述γ_1, γ_2的某一可选取数值，此时$\theta_1 = \frac{2}{3}$，$\theta_2 = \frac{1}{3}$，

而此时γ_1,γ_2的最优取值应为$\gamma_2=0$，$\gamma_1\in(0,1)$，不满足上述情形 a，故在情形 a 不存在纳什均衡。

当θ_1,θ_2满足：

$\theta_1=0$，$\theta_2\in(0,1)$

仅当$\theta_1=0$，$\theta_2=\frac{2}{3}$时，θ_1,θ_2取类似于上述γ_1,γ_2的某一可选取数值，而此时γ_1,γ_2的最优取值应为

$\gamma_2=1-\gamma_1$，$\gamma_1\in(0,1)$

不满足上述情形 b，故在情形 b 不存在纳什均衡。

当θ_1,θ_2满足：

$\theta_2=0$，$\theta_1\in(0,1)$

仅当$\theta_2=0$，$\theta_1=\frac{1}{3}$时，θ_1,θ_2取类似于上述γ_1,γ_2的某一可选取数值，而此时γ_1,γ_2的最优取值应为

$\gamma_1=0$，$\gamma_2\in(0,1)$

不满足上述情形 c。

当θ_1,θ_2满足：

$\theta_1\neq 1$，$\theta_2\neq 1$，$\theta_1+\theta_2>0$

当θ_1,θ_2位于上述情形 a 规定的γ_1,γ_2范围内时，最优的γ_1,γ_2选择位于$\gamma_2=1-\gamma_1$，$\gamma_1\in(0,1)$，不满足上述情形 d；当θ_1,θ_2位于上述情形 b 规定的γ_1,γ_2范围时，最优的γ_1,γ_2选择位于$\gamma_1=0$，$\gamma_2\in(0,1)$，不满足上述情形 d；当θ_1,θ_2位于上述情形 c 规定的γ_1,γ_2范围时，最优的γ_1,γ_2选择位于$\gamma_2=0$，$\gamma_1\in(0,1)$，不满足上述情形 d；最后，当θ_1,θ_2位于上述情形 d 规定的γ_1,γ_2范围时，最优的γ_1,γ_2位于$\gamma_1\neq 1$，$\gamma_2\neq 1$，$\gamma_1+\gamma_2>0$，满足上述情形 d；于是，仅有的纳什均衡为$\left\{\left(\frac{1}{3},\frac{1}{3},\frac{1}{3}\right),\left(\frac{1}{3},\frac{1}{3},\frac{1}{3}\right)\right\}$。

通过上例，我们可归纳出一种求解两人混合战略博弈纳什均衡的一般性方法。

一般地，设两人战略式表述博弈$G=(S_1,S_2;u_1,u_2)$，其中纯战略空间$S_1=\{s_{1h}\}_{h\in H}$，$S_2=\{s_{2p}\}_{p\in P}$，其中H和P为指标集。

设构成两人混合战略博弈纳什均衡的均衡战略组合为$\sigma=(\sigma_1,\sigma_2)$，其中$\sigma_1=\{\theta_{1h}\}_{h\in H}$，$\sigma_2=\{\gamma_{2p}\}_{p\in P}$，$0\le\theta_{1h}\le 1$，$0\le\gamma_{2p}\le 1$，$h\in H$，$p\in P$，$\sum_{h\in H}\theta_{1h}=1$，$\sum_{p\in P}\gamma_{2p}=1$，我们有

$$v_1=\sum_{h\in H}\theta_{1h}\sum_{p\in P}\gamma_{2p}u_1(s_{1h},s_{2p})=\sum_{h\in H}\theta_{1h}\overline{v_{1h}} \qquad (3.10)$$

其中$\overline{v_{1h}}=\sum_{p\in P}\gamma_{2p}u_1(s_{1h},s_{2p})$

给定σ_2，局中人 1 选取最优的σ_1使其v_1最大化，因为我们这里考虑的是混合战略而非纯战略均衡，若记

$$M_1 = \max_{h\in H}\overline{\gamma_{1h}}$$

（若M_1存在。在有限博弈中显然M_1存在），则必至少存在$h_1 \neq h_2$，$h_1, h_2 \in H$，使$\bar{v}_{1h1} = \bar{v}_{1h2} = M_1$。

记由H的所有非单元集子集构成的集合为$2^H - H$，即

$$2^H - H = \{a \mid a \text{ 为 } H \text{ 的子集，} a \neq \{h\}_{h\in H}\}$$

任给$a \in 2^H - H$，由下列条件决定$\sum_2$中的一个子集$\sum_{2a} \subset \sum_2$，

$$\bar{v}_{1h_1} = \bar{v}_{1h_2} = M_1,\ \forall h_1,\ h_2 \in a$$

$$\bar{v}_{1h} < M_1,\ \forall h \notin a,\ h \in H \qquad (3.11)$$

同样，我们有

$$v_2 = \sum_{p\in P}\gamma_{2p}\sum_{h\in H}\theta_{1h}u_2(s_{1h}, s_{2p}) = \sum_{p\in P}\gamma_{2p}\bar{v}_{2p} \qquad (3.12)$$

其中$\bar{v}_{2p} = \sum_{h\in H}\theta_{1h}u_2(s_{1h}, s_{2p})$

给定σ_1，局中人 2 选取最优的σ_2使其v_2最大化，因为我们这里考虑的是混合战略而非纯战略均衡，若记

$$M_2 = \sum_{p\in P}\bar{v}_{2p}$$

（若M_2存在。在有限博弈中显然M_2存在），则必至少存在$p_1 \neq p_2$，$p_1, p_2 \in P$，使$\bar{v}_{2p_1} = \bar{v}_{2p_2} = M_2$，记由$P$的所有非单元集子集构成的集合为$2^P - P$

即$2^P - P = \{b \mid b \text{ 为 } P \text{ 的子集，} b \neq \{p\}_{p\in P}\}$

任给$b \in 2^P - P$，由下列条件决定$\sum_1$中的一个子集$\sum_{1b} \subset \sum_1$，

$$\bar{v}_{2p_1} = \bar{v}_{2p_2} = M_2,\ \forall p_1, p_2 \in b$$

$$\bar{v}_{2p} < M_2,\ \forall p \notin b,\ p \in P \qquad (3.13)$$

显然，给定$\sigma_2 \in \sum_{2a}$，局中人 1 的最优混合战略由下列条件决定：

$$\sum_{h\in a}\theta_{1h} = 1,\ \theta_{1h} = 0, h \notin a,\ 0 \le \theta_{1h} < 1,\ h \in a \qquad (3.14)$$

记由条件（3.13）决定的$\sigma_1 \in \sum_1$构成的$\sum_1$中的子集为$\sum_{1a} \subset \sum_1$，易见的是$\sum_{1a}$与σ_2无关，即对于$\sum_{2a}$中的任何元素$\sigma_2 \in \sum_{2a}$，由条件（3.13）决定的$\sum_1$中的子集$\sum_{1a}$都是相同的，故$\sum_{1a}$只与a有关（因a通过（3.11）决定了$\sum_{2a}$）。

同样，给定$\sigma_1 \in \sum_{1b}$，局中人 2 的最优混合战略由下列条件决定：

$$\sum_{p\in b}\gamma_{2p} = 1,\ \gamma_{2p} = 0,\ p \notin b,\ 0 \le \gamma_{2p} < 1,\ p \in b \qquad (3.15)$$

记由条件（3.15）决定的$\sigma_2 \in \sum_2$构成的$\sum_2$中的子集为$\sum_{2b} \subset \sum_2$，易见的是$\sum_{2b}$与σ_1无关，即对于$\sum_{1b}$中的任何无素$\sigma_1 \in \sum_{1b}$，由条件（3.15）决定的$\sum_2$中的子集$\sum_{2b}$都是相同的，故$\sum_{2b}$只与b有关（因b通过（3.13）决定了$\sum_{1b}$）。

令
$$\sum_{1ab}=\sum_{1a}\cap\sum_{1b}$$
$$\sum_{2ab}=\sum_{2a}\cap\sum_{2b} \quad (3.16)$$

$\forall\sigma_1\in\sum_{1ab}$，$\forall\sigma_2\in\sum_{2ab}$，则$\sigma=(\sigma_1,\sigma_2)$为一混合战略纳什均衡。进一步，我们有如下定理。

定理 3.3 （蒲勇健，2001）$\sigma=(\sigma_1,\sigma_2)$是一混合博弈纳什均衡当且仅当$\sigma_1\in\sum_{1ab}$，$\sigma_2\in\sum_{2ab}$，$a\in 2^H-H$，$b\in 2^P-P$

证明：$\forall\sigma_1\in\sum_{1ab}$，$\forall\sigma_2\in\sum_{2ab}$，因$\sigma_1\in\sum_{1a}$，故$\sigma_1$是局中人 1 在给定$\sum_{2a}$中元素下的最优战略，因$\sigma_2\in\sum_{2a}$，故当然$\sigma_1$是局中人 1 在给定$\sigma_2$下的最优战略；同样，因$\sigma_2\in\sum_{2b}$，故$\sigma_2$是局中人 2 在给定$\sum_{1b}$中元素下的最优战略；因$\sigma_1\in\sum_{1b}$，当然$\sigma_2$是局中人 2 在给定$\sigma_1$下的最优战略。所以，$\sigma=(\sigma_1,\sigma_2)$是一个纳什均衡。

设$\sigma=(\sigma_1,\sigma_2)$为一个混合战略纳什均衡，则由推论 3.2.1 知必存在某$a\in 2^H-H$，$b\in 2^P-P$使

$\sigma_1\in\sum_{1b}$，$\sigma_2\in\sum_{2a}$

因σ_2是给定局中人 1 选择σ_1下局中人 2 的最优战略，故$\sigma_2\in\sum_{2b}$，所以$\sigma_2\in\sum_{2a}\cap\sum_{2b}=\sum_{2ab}$，同理有$\sigma_1\in\sum_{1ab}$。

证毕！

下面，我们运用定理 3.3 给出的方法来找出一些混合博弈的所有纳什均衡解。

例 3.17 “棍子·老虎·鸡·虫子”博弈的纳什均衡求解。

根据例 3.11 和表 3.13，我们知道该博弈也是一个对称博弈。

显然可猜测有一个对称均衡为$\left\{\left(\frac{1}{4},\frac{1}{4},\frac{1}{4},\frac{1}{4}\right),\left(\frac{1}{4},\frac{1}{4},\frac{1}{4},\frac{1}{4}\right)\right\}$，但是，我们还不知道该博弈除了这个对称均衡外是否还有其他均衡。下面，运用定理 3.3 提供的方法，我们来找出该博弈的全部纳什均衡。

显然，此例中可令$H=\{1,2,3,4\}$，$P=\{1,2,3,4\}$

令$\sigma_1=(\theta_1,\theta_2,\theta_3,1-\theta_1-\theta_2-\theta_3)$，$\sigma_2=(\gamma_1,\gamma_2,\gamma_3,1-\gamma_1-\gamma_2-\gamma_3)$，其中$0\le\theta_h<1$，$h=1,2,3$，$0\le\gamma_p<1$，$p=1,2,3$，$\sum_{h=1}^{3}\theta_h\le 1$，$\sum_{p=1}^{3}\gamma_p\le 1$

为了求出$\sum_{1a},\sum_{1b},\sum_{2a},\sum_{2b}$，需要求解以下几个联立方程（不等式）组：

$$\begin{cases}\gamma_2+(-1)(1-\gamma_1-\gamma_2-\gamma_3)=(-1)\gamma_1+\gamma_3\\-\gamma_2+(1-\gamma_1-\gamma_2-\gamma_3)<(-1)\gamma_1+\gamma_3\\\gamma_1-\gamma_3<(-1)\gamma_1+\gamma_3\end{cases} \quad (3.17)$$

$$\begin{cases}\gamma_2+(-1)\cdot(1-\gamma_1-\gamma_2-\gamma_3)=(-1)\cdot\gamma_2+1-\gamma_1-\gamma_2-\gamma_3\\-\gamma_1+\gamma_3<\gamma_2+(-1)\cdot(1-\gamma_1-\gamma_2-\gamma_3)\\\gamma_1-\gamma_3<\gamma_2+(-1)\cdot(1-\gamma_1-\gamma_2-\gamma_3)\end{cases} \quad (3.18)$$

$$\begin{cases}\gamma_2+(-1)(1-\gamma_1-\gamma_2-\gamma_3)=\gamma_1-\gamma_3\\-\gamma_1+\gamma_3<\gamma_1-\gamma_3\\-\gamma_2+(1-\gamma_1-\gamma_2-\gamma_3)<\gamma_1-\gamma_3\end{cases}\tag{3.19}$$

$$\begin{cases}-\gamma_1+\gamma_3=-\gamma_2+(1-\gamma_1-\gamma_2-\gamma_3)\\\gamma_2+(-1)(1-\gamma_1-\gamma_2-\gamma_3)<-\gamma_1+\gamma_3\\\gamma_1-\gamma_3<-\gamma_1+\gamma_3\end{cases}\tag{3.20}$$

$$\begin{cases}-\gamma_1+\gamma_3=\gamma_1-\gamma_3\\\gamma_2+(-1)(1-\gamma_1-\gamma_2-\gamma_3)<-\gamma_1+\gamma_3\\-\gamma_2+(1-\gamma_1-\gamma_2-\gamma_3)<-\gamma_1+\gamma_3\end{cases}\tag{3.21}$$

$$\begin{cases}-\gamma_2+(1-\gamma_1-\gamma_2-\gamma_3)=\gamma_1-\gamma_3\\\gamma_2+(-1)(1-\gamma_1-\gamma_2-\gamma_3)<\gamma_1-\gamma_3\\-\gamma_1+\gamma_3<\gamma_1-\gamma_3\end{cases}\tag{3.22}$$

$$\begin{cases}\gamma_2+(-1)(1-\gamma_1-\gamma_2-\gamma_3)=-\gamma_1+\gamma_3\\\gamma_2+(-1)(1-\gamma_1-\gamma_2-\gamma_3)=-\gamma_2+(1-\gamma_1-\gamma_2-\gamma_3)\\\gamma_1-\gamma_3<-\gamma_1+\gamma_3\end{cases}\tag{3.23}$$

$$\begin{cases}\gamma_2+(-1)(1-\gamma_1-\gamma_2-\gamma_3)=-\gamma_1+\gamma_3\\\gamma_2+(-1)(1-\gamma_1-\gamma_2-\gamma_3)=\gamma_1-\gamma_3\\-\gamma_2+(1-\gamma_1-\gamma_2-\gamma_3)<\gamma_1-\gamma_3\end{cases}\tag{3.24}$$

$$\begin{cases}\gamma_2+(-1)(1-\gamma_1-\gamma_2-\gamma_3)=-\gamma_2+(1-\gamma_1-\gamma_2-\gamma_3)\\\gamma_2+(-1)(1-\gamma_1-\gamma_2-\gamma_3)=\gamma_1-\gamma_3\\-\gamma_1+\gamma_3<\gamma_1-\gamma_3\end{cases}\tag{3.25}$$

$$\begin{cases}-\gamma_1+\gamma_3=-\gamma_2+(1-\gamma_1-\gamma_2-\gamma_3)\\-\gamma_1+\gamma_3=\gamma_1-\gamma_3\\\gamma_2+(-1)(1-\gamma_1-\gamma_2-\gamma_3)<\gamma_1-\gamma_3\end{cases}\tag{3.26}$$

$$\begin{cases}\gamma_2+(-1)(1-\gamma_1-\gamma_2-\gamma_3)=-\gamma_1+\gamma_3\\-\gamma_1+\gamma_3=-\gamma_2+(1-\gamma_1-\gamma_2-\gamma_3)\\-\gamma_1+\gamma_3=\gamma_1-\gamma_3\end{cases}\tag{3.27}$$

解得

$$\Sigma_{2(1,2)}=\left\{\gamma_1,\gamma_2,\gamma_3,(1-\gamma_1-\gamma_2-\gamma_3)\middle|\begin{matrix}\gamma_1+\gamma_2=\frac{1}{2}\\\gamma_2+\gamma_3>\frac{1}{2}\\\gamma_1<\gamma_3\end{matrix}\right\}$$

$$\Sigma_{2(1,3)}=\left\{\gamma_1,\gamma_2,\gamma_3,(1-\gamma_1-\gamma_2-\gamma_3)\middle|\begin{array}{l}2\gamma_2+\gamma_1+\gamma_3=1\\ \gamma_1+\gamma_2>\frac{1}{2}\\ \gamma_2+\gamma_3>\frac{1}{2}\end{array}\right\}$$

$$\Sigma_{2(1,4)}=\left\{\gamma_1,\gamma_2,\gamma_3,(1-\gamma_1-\gamma_2-\gamma_3)\middle|\begin{array}{l}\gamma_2+\gamma_3=\frac{1}{2}\\ \gamma_1>\gamma_3\\ \gamma_1+\gamma_2>\frac{1}{2}\end{array}\right\}$$

$$\Sigma_{2(2,3)}=\left\{\gamma_1,\gamma_2,\gamma_3,(1-\gamma_1-\gamma_2-\gamma_3)\middle|\begin{array}{l}\gamma_2+\gamma_3=\frac{1}{2}\\ \gamma_1+\gamma_2<\frac{1}{2}\\ \gamma_1<\gamma_3\end{array}\right\}$$

$$\Sigma_{2(2,4)}=\left\{\gamma_1,\gamma_2,\gamma_3,(1-\gamma_1-\gamma_2-\gamma_3)\middle|\begin{array}{l}\gamma_1=\gamma_3\\ \gamma_1+\gamma_2<\frac{1}{2}\\ \gamma_2+\gamma_3>\frac{1}{2}\end{array}\right\}$$

$$\Sigma_{2(3,4)}=\left\{\gamma_1,\gamma_2,\gamma_3,(1-\gamma_1-\gamma_2-\gamma_3)\middle|\begin{array}{l}\gamma_1+\gamma_2=\frac{1}{2}\\ \gamma_2+\gamma_3<\frac{1}{2}\\ \gamma_1>\gamma_3\end{array}\right\}$$

$$\Sigma_{2(1,2,3)}=\left\{\gamma_1,\gamma_2,\gamma_3,(1-\gamma_1-\gamma_2-\gamma_3)\middle|\begin{array}{l}\gamma_1+\gamma_2=\frac{1}{2}\\ 2\gamma_2+\gamma_1+\gamma_3=1\\ \gamma_1<\gamma_3\end{array}\right\}$$

$$\sum_{2(1,2,4)} = \left\{ \gamma_1, \gamma_2, \gamma_3, (1-\gamma_1-\gamma_2-\gamma_3) \middle| \begin{array}{l} \gamma_1+\gamma_2 = \frac{1}{2} \\ \gamma_2+\gamma_3 = \frac{1}{2} \\ \gamma_1+\gamma_2 > \frac{1}{2} \end{array} \right\} = \phi$$

$$\sum_{2(1,3,4)} = \left\{ \gamma_1, \gamma_2, \gamma_3, (1-\gamma_1-\gamma_2-\gamma_3) \middle| \begin{array}{l} 2\gamma_2+\gamma_1+\gamma_3 = 1 \\ \gamma_2+\gamma_3 = \frac{1}{2} \\ \gamma_1 > \gamma_3 \end{array} \right\}$$

$$\sum_{2(2,3,4)} = \left\{ \gamma_1, \gamma_2, \gamma_3, (1-\gamma_1-\gamma_2-\gamma_3) \middle| \begin{array}{l} \gamma_2+\gamma_3 = \frac{1}{2} \\ \gamma_1 = \gamma_3 \\ \gamma_2+\gamma_3 < \frac{1}{2} \end{array} \right\}$$

$$\sum_{2(1,2,3,4)} = \left\{ \gamma_1, \gamma_2, \gamma_3, (1-\gamma_1-\gamma_2-\gamma_3) \middle| \begin{array}{l} \gamma_1+\gamma_2 = \frac{1}{2} \\ \gamma_2+\gamma_3 = \frac{1}{2} \\ \gamma_1 = \gamma_3 \end{array} \right\}$$

其中ϕ表示空集。

由对称性知有

$\sum_{1b} = \sum_{2a}$，当$a = b$，$a \in 2^H - H$，$b \in 2^P - P$[12]

给定$\sum_{2a}$，我们可以求出$\sum_{1a}$，结果如下：

$$\sum_{1a} |_{a=(1,2)} = \left\{ \theta_1, \theta_2, \theta_3, 1-\theta_1-\theta_2-\theta_3 \middle| \begin{array}{l} \theta_1+\theta_2 = 1 \\ \theta_3 = 0 \\ \theta_1 > 0 \\ \theta_2 > 0 \end{array} \right\}$$

$$\sum_{1a} |_{a=(1,3)} = \left\{ \theta_1, \theta_2, \theta_3, 1-\theta_1-\theta_2-\theta_3 \middle| \begin{array}{l} \theta_1+\theta_3 = 1 \\ \theta_2 = 0 \\ \theta_1 > 0 \\ \theta_3 > 0 \end{array} \right\}$$

$$\sum_{1a} |_{a=(1,4)} = \left\{ \theta_1, \theta_2, \theta_3, 1-\theta_1-\theta_2-\theta_3 \middle| \begin{array}{l} \theta_2 = 0 \\ \theta_3 = 0 \\ 0 < \theta_1 < 0 \end{array} \right\}$$

$$\Sigma_{1a}|_{a=(2,3)}=\left\{\theta_1,\theta_2,\theta_3,1-\theta_1-\theta_2-\theta_3\middle|\begin{array}{l}\theta_2+\theta_3=1\\ \theta_1=0\\ \theta_2>0\\ \theta_3>0\end{array}\right\}$$

$$\Sigma_{1a}|_{a=(2,4)}=\left\{\theta_1,\theta_2,\theta_3,1-\theta_1-\theta_2-\theta_3\middle|\begin{array}{l}\theta_1=0\\ \theta_3=0\\ 0<\theta_2<1\end{array}\right\}$$

$$\Sigma_{1a}|_{a=(3,4)}=\left\{\theta_1,\theta_2,\theta_3,1-\theta_1-\theta_2-\theta_3\middle|\begin{array}{l}\theta_1=0\\ \theta_2=0\\ 0<\theta_3<1\end{array}\right\}$$

$$\Sigma_{1a}|_{a=(1,2,3)}=\left\{\theta_1,\theta_2,\theta_3,1-\theta_1-\theta_2-\theta_3\middle|\begin{array}{l}\theta_1+\theta_2+\theta_3=1\\ 0\le\theta_1<1\\ 0\le\theta_2<1\\ 0\le\theta_3<1\end{array}\right\}$$

$$\Sigma_{1a}|_{a=(1,2,4)}=\left\{\theta_1,\theta_2,\theta_3,1-\theta_1-\theta_2-\theta_3\middle|\begin{array}{l}\theta_3=0\\ 0\le\theta_1<1\\ 0\le\theta_2<1\\ 0<\theta_1+\theta_2<1\end{array}\right\}$$

$$\Sigma_{1a}|_{a=(1,3,4)}=\left\{\theta_1,\theta_2,\theta_3,1-\theta_1-\theta_2-\theta_3\middle|\begin{array}{l}\theta_2=0\\ 0\le\theta_1<1\\ 0\le\theta_3<1\\ 0<\theta_1+\theta_2\le1\end{array}\right\}$$

$$\Sigma_{1a}|_{a=(2,3,4)}=\left\{\theta_1,\theta_2,\theta_3,1-\theta_1-\theta_2-\theta_3\middle|\begin{array}{l}\theta_1=0\\ 0\le\theta_2<1\\ 0\le\theta_3<1\\ 0<\theta_2+\theta_3\le1\end{array}\right\}$$

$$\Sigma_{1a}|_{a=(1,2,3,4)}=\left\{\theta_1,\theta_2,\theta_3,1-\theta_1-\theta_2-\theta_3\middle|\begin{array}{l}0<\theta_1+\theta_2+\theta_3<1\\ 0\le\theta_1<1\\ 0\le\theta_2<1\\ 0\le\theta_3<1\end{array}\right\}$$

根据对称性有$\Sigma_{2b}=\Sigma_{1a}$，当$a=b$，$a\in 2^H-H$，$b\in 2^P-P$。下面来计算Σ_{1ab}和Σ_{2ab}，直接给出结果。

$$\sum_{1(1,2)(1,2)}=\phi，\ \sum_{1(1,2)(1,3)}=\phi，\ \sum_{1(1,2)(1,4)}=\left\{\theta_1,\theta_2,\theta_3,1-\theta_1-\theta_2-\theta_3\left|\begin{aligned}&\theta_1=\frac{1}{2}\\&\theta_2=\frac{1}{2}\\&\theta_3=0\end{aligned}\right.\right\}$$

$\sum_{1(1,2)(2,3)}=\phi$， $\sum_{1(1,2)(2,4)}=\phi$， $\sum_{1(1,2)(3,4)}=\phi$， $\sum_{1(1,2)(1,2,3)}=\phi$

$\sum_{1(1,2)(1,2,4)}=\phi$， $\sum_{1(1,2)(1,3,4)}=\phi$， $\sum_{1(1,2)(2,3,4)}=\phi$， $\sum_{1(1,2)(1,2,3,4)}=\phi$

根据对称性，我们知

$\sum_{1(1,3)(1,2)}=\sum_{1(1,3)}\cap\sum_{1(1,2)}=\sum_{1(1,2)}\cap\sum_{1(1,3)}=\sum_{1(1,2)}\cap\sum_{2(1,3)}=\sum_{1(1,2)(1,3)}=\phi$

$\sum_{1(1,3)(1,3)}=\phi$， $\sum_{1(1,3)(1,4)}=\phi$， $\sum_{1(1,3)(2,3)}=\phi$， $\sum_{1(1,3)(2,4)}=\phi$， $\sum_{1(1,3)(3,4)}=\phi$，

$\sum_{1(1,3)(1,2,3)}=\phi$， $\sum_{1(1,3)(1,2,4)}=\phi$， $\sum_{1(1,3)(1,3,4)}=\phi$， $\sum_{1(1,3)(2,3,4)}=\phi$，

$$\sum_{1(1,3)(1,2,3,4)}=\left\{\theta_1,\theta_2,\theta_3,1-\theta_1-\theta_2-\theta_3\left|\begin{aligned}&\theta_1=\frac{1}{2}\\&\theta_2=0\\&\theta_3=\frac{1}{2}\end{aligned}\right.\right\};$$

$\sum_{1(1,4)(1,2)}=\phi$， $\sum_{1(1,4)(1,3)}=\phi$， $\sum_{1(1,4)(1,4)}=\phi$， $\sum_{1(1,4)(2,3)}=\phi$， $\sum_{1(1,4)(2,4)}=\phi$，

$$\sum_{1(1,4)(3,4)}=\left\{\theta_1,\theta_2,\theta_3,1-\theta_1-\theta_2-\theta_3\left|\begin{aligned}&\theta_1=\frac{1}{2}\\&\theta_2=0\\&\theta_3=0\end{aligned}\right.\right\}，\ \sum_{1(1,4)(1,2,3)}=\phi，\ \sum_{1(1,4)(1,2,4)}=\phi，$$

$\sum_{1(1,4)(1,3,4)}=\phi$， $\sum_{1(1,4)(2,3,4)}=\phi$， $\sum_{1(1,4)(1,2,3,4)}=\phi$；

$$\sum_{1(2,3)(1,2)}=\left\{\theta_1,\theta_2,\theta_3,1-\theta_1-\theta_2-\theta_3\left|\begin{aligned}&\theta_1=0\\&\theta_2=\frac{1}{2}\\&\theta_3=\frac{1}{2}\end{aligned}\right.\right\}，\ \sum_{1(2,3)(1,3)}=\phi，\ \sum_{1(2,3)(1,4)}=\phi，$$

$\sum_{1(2,3)(2,3)}=\phi$， $\sum_{1(2,3)(2,4)}=\phi$， $\sum_{1(2,3)(3,4)}=\phi$， $\sum_{1(2,3)(1,2,3)}=\phi$， $\sum_{1(2,3)(1,2,4)}=\phi$，

$\sum_{1(2,3)(1,3,4)}=\phi$， $\sum_{1(2,3)(2,3,4)}=\phi$， $\sum_{1(2,3)(1,2,3,4)}=\phi$；

$\sum_{1(2,4)(1,2)}=\phi$， $\sum_{1(2,4)(1,3)}=\phi$， $\sum_{1(2,4)(1,4)}=\phi$， $\sum_{1(2,4)(2,3)}=\phi$， $\sum_{1(2,4)(2,4)}=\phi$，

$\sum_{1(2,4)(3,4)}=\phi$， $\sum_{1(2,4)(1,2,3)}=\phi$， $\sum_{1(2,4)(1,2,4)}=\phi$， $\sum_{1(2,4)(1,3,4)}=\phi$，

$$\Sigma_{1(2,4)(2,3,4)}=\phi\text{，}\quad \Sigma_{1(2,4)(1,2,3,4)}=\left\{\theta_1,\theta_2,\theta_3,1-\theta_1-\theta_2-\theta_3 \middle| \begin{matrix}\theta_1=0\\ \theta_2=\frac{1}{2}\\ \theta_3=0\end{matrix}\right\}\text{；}$$

$$\Sigma_{1(3,4)(1,2)}=\phi\text{，}\quad \Sigma_{1(3,4)(1,3)}=\phi\text{，}\quad \Sigma_{1(3,4)(1,4)}=\phi\text{，}$$

$$\Sigma_{1(3,4)(2,3)}=\left\{\theta_1,\theta_2,\theta_3,1-\theta_1-\theta_2-\theta_3 \middle| \begin{matrix}\theta_1=0\\ \theta_2=0\\ \theta_3=\frac{1}{2}\end{matrix}\right\}\text{，}\quad \Sigma_{1(3,4)(2,4)}=\phi\text{，}\quad \Sigma_{1(3,4)(3,4)}=\phi\text{，}$$

$$\Sigma_{1(3,4)(1,2,3)}=\phi\text{，}\quad \Sigma_{1(3,4)(1,2,4)}=\phi\text{，}\quad \Sigma_{1(3,4)(1,3,4)}=\phi\text{，}\quad \Sigma_{1(3,4)(2,3,4)}=\phi\text{，}$$

$$\Sigma_{1(3,4)(1,2,3,4)}=\phi\text{；}$$

$$\Sigma_{1(1,2,3)(1,2)}=\left\{\theta_1,\theta_2,\theta_3,1-\theta_1-\theta_2-\theta_3 \middle| \begin{matrix}0\le\theta_1=\frac{1}{2}\\ 0<\theta_2<1\\ \theta_3=\frac{1}{2}\end{matrix}\right\}$$

$$\Sigma_{1(1,2,3)(1,3)}=\phi\text{，}\quad \Sigma_{1(1,2,3)(1,4)}=\left\{\theta_1,\theta_2,\theta_3,1-\theta_1-\theta_2-\theta_3 \middle| \begin{matrix}\theta_1=\frac{1}{2}\\ \theta_2>0\\ 0\le\theta_3<\frac{1}{2}\end{matrix}\right\}$$

$$\Sigma_{1(1,2,3)(2,3)}=\phi\text{，}\quad \Sigma_{1(1,2,3)(2,4)}=\left\{\theta_1,\theta_2,\theta_3,1-\theta_1-\theta_2-\theta_3 \middle| \begin{matrix}\theta_1=\frac{1}{2}\\ \theta_2=0\\ \theta_3=\frac{1}{2}\end{matrix}\right\}$$

$$\Sigma_{1(1,2,3)(3,4)}=\phi\text{，}\quad \Sigma_{1(1,2,3)(1,2,3)}=\phi\text{，}\quad \Sigma_{1(1,2,3)(1,2,4)}=\phi\text{，}\quad \Sigma_{1(1,2,3)(1,3,4)}=\phi\text{，}$$

$$\Sigma_{1(1,2,3)(2,3,4)}=\phi\text{，}\quad \Sigma_{1(1,2,3)(1,2,3,4)}=\left\{\theta_1,\theta_2,\theta_3,1-\theta_1-\theta_2-\theta_3 \middle| \begin{matrix}\theta_1=\frac{1}{2}\\ \theta_2=0\\ \theta_3=\frac{1}{2}\end{matrix}\right\}\text{；}$$

$$\Sigma_{1(1,2,4)(1,2)}=\phi\text{，}\quad \Sigma_{1(1,2,4)(1,3)}=\left\{\theta_1,\theta_2,\theta_3,1-\theta_1-\theta_2-\theta_3 \middle| \begin{matrix}0<\theta_1<\frac{1}{2}\\ \frac{1}{2}<\theta_2<1-\theta\\ \theta_3=0\end{matrix}\right\}\text{，}$$

$$\sum\nolimits_{1(1,2,4)(1,4)} = \left\{ \theta_1, \theta_2, \theta_3, 1-\theta_1-\theta_2-\theta_3 \middle| \begin{array}{l} 0 < \theta_1 < \frac{1}{2} \\ \theta_2 = \frac{1}{2} \\ \theta_3 = 0 \end{array} \right\}, \quad \sum\nolimits_{1(1,2,4)(2,3)} = \phi,$$

$$\sum\nolimits_{1(1,2,4)(2,4)} = \phi, \quad \sum\nolimits_{1(1,2,4)(3,4)} = \left\{ \theta_1, \theta_2, \theta_3, 1-\theta_1-\theta_2-\theta_3 \middle| \begin{array}{l} 0 < \theta_1 < 1 \\ \theta_1 + \theta_2 = \frac{1}{2} \\ \theta_3 = 0 \end{array} \right\},$$

$\sum_{1(1,2,4)(1,2,3)} = \phi$， $\sum_{1(1,2,4)(1,2,4)} = \phi$， $\sum_{1(1,2,4)(1,3,4)} = \phi$， $\sum_{1(1,2,4)(2,3,4)} = \phi$，

$$\sum\nolimits_{1(1,2,4)(1,2,3,4)} = \left\{ \theta_1, \theta_2, \theta_3, 1-\theta_1-\theta_2-\theta_3 \middle| \begin{array}{l} \theta_1 = 0 \\ \theta_2 = \frac{1}{2} \\ \theta_3 = 0 \end{array} \right\};$$

$\sum_{1(1,3,4)(1,2)} = \phi$， $\sum_{1(1,3,4)(1,3)} = \phi$， $\sum_{1(1,3,4)(1,4)} = \phi$，

$$\sum\nolimits_{1(1,3,4)(2,3)} = \left\{ \theta_1, \theta_2, \theta_3, 1-\theta_1-\theta_2-\theta_3 \middle| \begin{array}{l} \theta_1 < \frac{1}{2} \\ \theta_2 = 0 \\ \theta_3 = \frac{1}{2} \end{array} \right\},$$

$$\sum\nolimits_{1(1,3,4)(3,4)} = \left\{ \theta_1, \theta_2, \theta_3, 1-\theta_1-\theta_2-\theta_3 \middle| \begin{array}{l} \theta_1 = \frac{1}{2} \\ \theta_2 = 0 \\ \theta_3 < \frac{1}{2} \end{array} \right\};$$

$\sum_{1(1,3,4)(1,2,3)} = \phi$， $\sum_{1(1,3,4)(1,2,4)} = \phi$， $\sum_{1(1,3,4)(1,3,4)} = \phi$， $\sum_{1(1,3,4)(2,3,4)} = \phi$，

$$\sum\nolimits_{1(1,3,4)(1,2,3,4)} = \left\{ \theta_1, \theta_2, \theta_3, 1-\theta_1-\theta_2-\theta_3 \middle| \begin{array}{l} \theta_1 = \frac{1}{2} \\ \theta_2 = 0 \\ \theta_3 = \frac{1}{2} \end{array} \right\};$$

$\sum_{1(2,3,4)(1,2)} = \phi$， $\sum_{1(2,3,4)(1,3)} = \phi$， $\sum_{1(2,3,4)(1,4)} = \phi$，

$$\sum\nolimits_{1(2,3,4)(2,3)} = \left\{ \theta_1, \theta_2, \theta_3, 1-\theta_1-\theta_2-\theta_3 \left| \begin{array}{l} \theta_1 = 0 \\ 0 \le \theta_2 < \dfrac{1}{2} \\ \theta_2 + \theta_3 = \dfrac{1}{2} \end{array} \right. \right\}$$

$\sum_{1(2,3,4)(2,4)} = \phi$， $\sum_{1(2,3,4)(3,4)} = \phi$， $\sum_{1(2,3,4)(1,2,3)} = \phi$， $\sum_{1(2,3,4)(1,2,4)} = \phi$，

$\sum_{1(2,3,4)(1,3,4)} = \phi$， $\sum_{1(2,3,4)(2,3,4)} = \phi$，

$$\sum\nolimits_{1(2,3,4)(1,2,3,4)} = \left\{ \theta_1, \theta_2, \theta_3, 1-\theta_1-\theta_2-\theta_3 \left| \begin{array}{l} \theta_1 = 0 \\ \theta_2 = \dfrac{1}{2} \\ \theta_3 = 0 \end{array} \right. \right\}$$

显然，因$\sum_{1a} |_{a=(1,2,3,4)} \supseteq \sum_{1b}$

故

$\sum_{1(1,2,3,4)b} = \sum_{1b}$， $b \in 2^P - p$

因为由对称性知

$\sum_{1ab} = \sum_{2ab}$

故毋须计算$\sum_{2ab}$。

从以上计算结果看，除$\sum_{1(1,2,3,4)(1,2,3,4)}$和$\sum_{2(1,2,3,4)(1,2,3,4)}$外，还有许多$\sum_{1ab}$（和$\sum_{2ab}$）是非空集，因此，该博弈除了对称均衡外，还存在其它无数多个均衡，特别地当$a = b = (1,2,3,4)$时，我们得到$\sum_{1(1,2,3,4)(1,2,3,4)} = \sum_{2(1,2,3,4)(1,2,3,4)} \ni \left(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right)$，这就是对称均衡。这种计算混合战略博弈均衡的方法看起来相当繁锁，但却十分适于计算机运算，因为推算过程是十分程式化的。

例 3.18　田忌赛马博弈

我们在 3.2.3 小节中已介绍了这个博弈的历史故事内容，在这里，我们假设田忌（和孙膑）与齐威王在智力上是不分高下的（或许违反了历史真实），并且在这里会自然想到该博弈不存在纯战略纳什均衡。运用定理 3.3 给出的方法，我们可以求出这个博弈的所有混合战略博弈纳什均衡解，但这可能面临十分繁锁的数学计算，我们在这里不打算展开这种计算，只给出其中一个对称均衡，读者可以自行去求解其它的均衡。

齐威王和田忌的纯战略空间都是由各自的出马顺序所构成，即：

$S_{齐威王}$={上中下，上下中，中上下，中下上，下上中，下中上}

$S_{田忌}$={上中下，上下中，中上下，中下上，下上中，下中上}

表 3.17 给出了这个博弈的支付矩阵。

表 3.17　田忌赛马博弈

		田忌					
		上中下	上下中	中上下	中下上	下上中	下中上
齐威王	上中下	3，–3	1，–1	1，–1	1，–1	–1，1	1，–1
	上下中	1，–1	3，–3	1，–1	1，–1	1，–1	–1，1
	中上下	1，–1	–1，1	3，–3	1，–1	1，–1	1，–1
	中下上	–1，1	1，–1	1，–1	3，–3	1，–1	1，–1
	下上中	1，–1	1，–1	1，–1	–1，1	3，–3	1，–1
	下中上	1，–1	1，–1	–1，1	1，–1	1，–1	3，–3

设齐威王和田忌的均衡混合战略分别为：$\sigma_1=\left(\theta_1,\theta_2,\theta_3,\theta_4,\theta_5,1-\sum_{k=1}^{5}\theta_k\right)$和$\sigma_2=\left(\gamma_1,\gamma_2,\gamma_3,\gamma_4,\gamma_5,1-\sum_{k=1}^{5}\gamma_k\right)$，其中$0\le\theta_k<1$，$0\le\gamma_k<1$，$k=1,\cdots,5$，$1>\sum_{k=1}^{5}\theta_k>0$，$1>\sum_{k=1}^{5}\gamma_k>0$，据推论 3.2.1 有：

$$\begin{cases}3\gamma_1+\gamma_2+\gamma_3+\gamma_4-\gamma_5+\left(1-\sum_{k=1}^{5}\gamma_k\right)=\gamma_1+3\gamma_2+\gamma_3+\gamma_4+\gamma_5-\left(1-\sum_{k=1}^{5}\gamma_k\right)\\3\gamma_1+\gamma_2+\gamma_3+\gamma_4-\gamma_5+\left(1-\sum_{k=1}^{5}\gamma_k\right)=\gamma_1-\gamma_2+3\gamma_3+\gamma_4+\gamma_5-\left(1-\sum_{k=1}^{5}\gamma_k\right)\\3\gamma_1+\gamma_2+\gamma_3+\gamma_4-\gamma_5+\left(1-\sum_{k=1}^{5}\gamma_k\right)=-\gamma_1+\gamma_2+\gamma_3+3\gamma_4+\gamma_5+\left(1-\sum_{k=1}^{5}\gamma_k\right)\\3\gamma_1+\gamma_2+\gamma_3+\gamma_4-\gamma_5+\left(1-\sum_{k=1}^{5}\gamma_k\right)=\gamma_1+\gamma_2+\gamma_3-\gamma_4+3\gamma_5+\left(1-\sum_{k=1}^{5}\gamma_k\right)\\3\gamma_1+\gamma_2+\gamma_3+\gamma_4-\gamma_5+\left(1-\sum_{k=1}^{5}\gamma_k\right)=\gamma_1+\gamma_2-\gamma_3+\gamma_4+\gamma_5+3\left(1-\sum_{k=1}^{5}\gamma_k\right)\end{cases}\tag{3.28}$$

解得：$\gamma_1=\gamma_2=\gamma_3=\gamma_4=\gamma_5=\frac{1}{6}$，据对称性有$\theta_1=\theta_2=\theta_3=\theta_4=\theta_5=\frac{1}{6}$，故均衡为$\left\{\left(\frac{1}{6},\frac{1}{6},\frac{1}{6},\frac{1}{6},\frac{1}{6},\frac{1}{6}\right),\left(\frac{1}{6},\frac{1}{6},\frac{1}{6},\frac{1}{6},\frac{1}{6},\frac{1}{6}\right)\right\}$。

据推论 3.2.1，将此结果代入式（3.28）中任一等式的任一端，得齐威王的期望支付为 1，同样可得田忌的期望支付为–1。所以，如果我们不假定齐威王是“傻子”，则在此均衡中齐威王平均看来在一次赛马中会赢去田忌一千斤铜，这是因为齐威王的三匹马略胜田忌的三匹马一筹，从而齐威王具有绝对优势的缘故。

例 3.19　华容道博弈的纳什均衡

赤壁大战后，曹操率残兵败将逃命，孔明派关羽领兵阻截曹兵，曹操有两条路可走，一条路是华容道，除此之外还有另外一条路可逃，关羽也可任选一条路设伏兵捉拿曹操，与《三国演义》所述不同，我们这里假定孔明与曹操之间也不存在智力上的差别。表 3.17 给出这个博弈的支付矩阵。因为关羽与孔明曾立下军令状——逮不住曹操回去后会受到惩罚，故关羽在逮不着曹操情况下的预期支付为负（事实上，关羽放走了曹操，回去后经刘备特赦未受到惩罚）。

表 3.18　华容道博弈

		曹操	
		华容道	另一条路
关	华容道	1，–1	–1，1
羽	另一条路	–1，1	1，–1

不难解出该博弈的均衡只有一个，即对称均衡$\left\{\left(\frac{1}{2},\frac{1}{2}\right),\left(\frac{1}{2},\frac{1}{2}\right)\right\}$，曹操和关羽的期望支付都为零，即曹操有一半的机会逃脱，而关羽也只有一半的可能捉住曹操，因此，除非我们假定曹操是“傻子”（如《三国演义》中所写的那样），关羽（孔明）不一定能逮住曹操。

例 3.20　监督博弈

正如我们在第二章中所指出的那样，许多不同的经济抑或社会现象都可以按博弈模型归类，一些看来是完全不同的现象却在博弈论专家看来是没有什么本质区别的。我们这里介绍的监督博弈实际上刻画了诸如税收检查，质量检查，惩治犯罪，雇主监督雇员等许多现象的博弈论本质。所谓监督博弈来自于监督检查存在成本，因而监督者不会总是对被监督者的所有情形都实施检查，而是随机地采取检查或不检查的策略，被监督者也知道监督者的这种策略选择，因而也以随机的方式选择努力或不努力的策略。这样，监督者与被监督者之间就展开混合战略博弈，我们现在感兴趣的是找出这种博弈的均衡。下面以税收检查为例来说明之。这个博弈的局中人是税收机关和纳税人。税收机关的纯战略空间中包括“检查”和“不检查”两个元素，纳税人的纯战略空间中包括“逃税”和“不逃税”两个元素。表 3.19 给出了这个博弈的支付矩阵，其中，a 为应纳税款，c 是检查成本，F 是罚款量。为使模型有意义，假定 $C<a+F$。此时不存在纯战略纳什均衡。

表 3.19　监督博弈

		纳税人	
		逃税	不逃税
税收机关	检查	a–c+F, –a–F	a–c, –a
	不检查	0, 0	a, –a

设税收机关检查的概率为θ，纳税人逃税的概率为γ。据推论 3.2.1 不难求得均衡为$\left\{\left(\frac{a}{a+F},\frac{F}{a+F}\right),\left(\frac{c}{a+F},\frac{a+F-c}{a+F}\right)\right\}$，显然，对逃税的惩罚愈重，应纳税款愈多，纳税人逃税的概率就愈小；检查成本愈高，纳税人逃税的概率就愈大。为什么应纳税款愈多，纳税人逃税的概念反而愈小呢？这是因为，应纳税款愈多，税收机关检查的概率就愈高，逃税被抓住的可能性就愈大，因而纳税人反而不敢逃税了。这一点或许可以解释为什么逃税现象在小企业中比在大企业中更为普遍，在低收入阶层比在高收入阶层更普遍。[13]

3.2.11 奇数定理

某些博弈可能既存在纯战略纳什均衡，又存在混合战略博弈纳什均衡。表 3.20 给出一个被称为“性别战”（battle of the sexes）的博弈，这个博弈说的是一对处于热恋中的青年男女准备来一次约会，男孩喜欢欣赏足球比赛而女孩喜欢看芭蕾舞表演。当然，他们俩都更喜欢在一起看表演而不愿分离，分离（哪怕是短暂的）带给他俩的痛苦会完全抵消掉独自看自己喜欢的表演所带来的快乐，当心爱的人儿不在身边时，哪有心思去欣赏表演或比赛呢？

表 3.20　性别战博弈

		女	
		足球	芭蕾
男	足球	2，1	0，0
	芭蕾	0，0	1，2

运用划线法，不难找到两个纯战略纳什均衡（足球，足球）和（芭蕾，芭蕾）。 我们再来看看该博弈是否还存在混合战略纳什均衡，据推论 3.2.1，我们可以很快找出混合战略纳什均衡为$\left\{\left(2/3,\frac{1}{3}\right),\left(\frac{1}{3},\frac{2}{3}\right)\right\}$。一般地，我们有如下定理：

定理 3.4(wilson,1971)　几乎所有有限博弈都有有限奇数个纳什均衡（包括混合战略均衡）。

该定理被称为奇数定理(oddness theorem)。定理中的“几乎所有”(almost all)是测度

论中的数学术语，指的是除掉一个测度为零的集合外，其他情形定理结论都成立。测度论是现代数学的一个分支，是古典积分学在现代的推广。这个定理说的是在“绝大多数”情况下，有限博弈都有有限奇数个纳什均衡。因此，如果一个有限博弈没有纯战略纳什均衡，则很可能有一个混合战略纳什均衡；如果一个有限博弈已发现有两个纯战略纳什均衡，则可能还存在一个混合战略纳什均衡。这个定理的证明过于数学化，这里就不再给出，对证明过程感兴趣的读者可参看 wllson (1971)。

例 3.21　斗鸡博弈

“斗鸡”是盛行于世界各地的一种古老的游戏，想必许多人对“斗鸡”中雄鸡那特有的姿势和进退动作有着深刻的印象；与其相类似的是奥林匹克运动员的击剑比赛（或许这种类比稍许欠雅，但本质上却是相同的)，击剑运动员们总是以这样一种战略进行比赛，即对方进攻时后退，对方后退时进攻，我们下面用博弈论模型证明无论是雄鸡还是击剑运动员，它(他)们的战略都是纳什均衡战略，同时，还将给出一种混合战略均衡，说明雄鸡（如果可能的话）或击剑运动员除采用上述战略外还可用一种混合战略进行攻击或后退。

表 3.21　斗鸡博弈

		2	
		进攻	后退
1	进攻	–3, –3	2, 0
	后退	0, 2	0, 0

表 3.21 给出斗鸡博弈的支付矩阵。

两个局中人都有两个纯战略：“进攻”和“后退”；若两个人都进攻，则两败俱伤；若一方进攻则另一方后退，进攻者取得胜利(或得分)，后退者丢了面子(或不得分)；若两人都后退则两人都丢面子(或都不得分)。

通过划线法找到两个纯战略纳什均衡，（进攻, 后退）和（后退, 进攻），据奇数定理，我们猜想还存在第三个纳什均衡，如果存在这个均衡，它必然是混合战略纳什增衡。据推论 3.2.1 不难找出这个均衡：$\left\{\left(\frac{2}{5},\frac{3}{5}\right),\left(\frac{2}{5},\frac{3}{5}\right)\right\}$。

这个例子可以解释许多观察到的现象，如某些公共品的供给。若村子里住的是两户富人，有一条路要修，一种可能是一家修路，另一家人就不会修；一家不修，另一家就得修。[14]

另外，在冷战期间，苏美两个军事集团在世界各地抢占地盘，也是一种斗鸡博弈。当一方已经抢占了一块地盘，另一方就设法占领另一块地盘，而不是与对手竞争同一块

地盘。我们还从电视上十分熟悉这样的镜头，当抗议者的游行队伍与防暴警察相接近时，必定有一方会主动后退。有时，夫妻间因事口角，当吵得厉害了，不是妻子回娘家呆几天，就是丈夫到院子里抽支烟。

3.2.12 关于混合战略博弈的解释

根据推论 3.2.1，我们知道混合战略博弈纳什均衡具有这样一个特点，即当局中人选择混合战略博弈纳什均衡战略时，他以正的概率所选择的那些纯战略给他带来的期望支付之间是无差异的。也就是说，当他选择某一个在混合战略博弈纳什均衡战略中是以正的概率选择的那些纯战略时，他预期获得的期望支付是一个常数。这样一来，我们就引出了一个深刻的问题：既然局中人选择那些在混合战略博弈纳什均衡战略中以正的概率选择的纯战略都带来同样的期望支付（根据推论 3.2.1，我们知道，这个期望支付还等于局中人选择混合战略博弈纳什均衡战略时所获得的期望支付），那么，存在什么理由使局中人实际上选择了某一纯战略而未能选择其它能带来同样期望支付的纯战略呢？既然局中人在这些纯战略之间是无差异的，那么，是什么样的依据使得局中人在它们之间进行选择呢？进一步，我们将纯战略也视为混合战略的一种特例，那些构成混合战略的纯战略与该混合战略本身一起构成一个混合战略的子集，那么，我们凭什么道理预测局中人会选择其中某个纯战略而没有选择其他纯战略呢？显然，这个问题十分类似于我们在前面所讨论的关于非严格纳什均衡存在的问题时所面临的困境，这也是我们在注释[11]中所谈到的尼尔斯·玻尔的“饿死的驴”所面临的问题——既然两堆草料完全相同，驴子凭什么选择先吃其中一堆草而后吃另一堆草。在尼尔斯·玻尔那里，他通过增加一个十分令人震惊的（所谓“革命性”的）假设来解决这一问题——假设驴子具有一种“自主性”或在粒子物理学中假设粒子的行为具有一种随机性，即所谓“上帝在掷骰子”。这样，驴子可以完全不依照机械的逻辑决定论来选择先吃其中一堆草，而这种选择完全是任意的或真正随机的。不需要任何因果决定关系，这样驴子才可能避免被活活饿死。当然，我们还可以从另一个角度来试图解决这一问题，即增添所谓“隐变量”假设。美国物理学家玻姆是爱因斯坦的忠实信徒，他也坚持认为上帝不会通过掷骰子的方式来安排宇宙的秩序。他提出了“隐变量”理论来解释粒子看似随机性的行为。

什么是“隐变量”理论呢？我们对物质的认识是按照我们所知的物质特征分类来进行的。譬如，现代物理学对基本粒子的了解无非是依据粒子的电荷量（包括电荷的正或负）、质量、自旋、运动速度及空间位置等特征。当我们发现两个粒子在上述特征方面相同时，我们会认为两个粒子是相同的。两个电子具有相同的质量、电荷、自旋，所以我们认为不同的电子都是相同的。当两个电子在同一窄孔处以同一速度穿过窄孔而打在衍射屏上时，我们发现具有同样特征的电子居然可以出现在衍射屏的不同位置上，这个实验结果显然与因果定律相矛盾。物理学家们不愿放弃因果律，因为它是整个现代科学

的基础。于是，海森堡提出了著名的“测不准原理”假设来替代放弃因果律的假设。根据“测不准原理”，我们不能同时无限精确地测定粒子的速度和位置，因而就不能说两个电子是在同一窄孔的同一位置以同一初始速度穿过窄孔，打在衍射屏上不同位置的电子可以是以不同的初始速度或不同的初始位置或两者同时存在的方式穿过窄孔的，因而当然可以出现在衍生屏上不同的位置上。量子物理学进一步确定了许多不能同时精确测定的指标对，即所谓“共轭物理量”，如速度与位置就是一对不可同时无限精确测定的“共轭物理量”。尼尔斯·玻尔是一位兼具哲学家禀质的物理学家，以他为首的哥本哈根学派认为基本粒子共轭物理量的测不准性质是宇宙固有的一种随机性质，基本粒子似乎具有“自由意志”，它的行为有一种不可预测性，但其随机变化的概率分布是可以预测的。例如，穿过窄孔的电子可以随机地打在衍射屏的不同位置上，预测电子具体会出现在衍射屏上的哪一个位置是徒劳的，但电子出现在衍射屏上不同位置的概率却是可以预测的（即通过所谓薛定谔方程中波函数计算）。根据哥本哈根学派的观点，粒子的测不准性质并不是由于人类观测自然现象手段或工具不足的制约，而是因为自然运动本身有一种固有的随机性或不可预测性。对于这种被爱因斯坦称为“上帝掷骰子”的理论，坚信决定论的其他物理学家不以为然。物理学家玻姆提出了“隐变量”理论与之抗衡，他认为我们对自然的认识过程是渐进的，由于人类认识能力和观测手段是一个不断发展和深化的过程，因而现阶段人类还不能对基本粒子除上述特征之外的其他更加细微精致的特征进行识别，有可能将不同的事物视为完全同一的东西。俗话说：“没有两片相同的树叶”，两个电子尽管具有相同的质量、电荷、自旋、位置和速度，但还有可能具有其他不同的特征使它们本质上属于不同的事物，因而尽管它们以相同的速度穿过窄孔中相同的位置，却不会出现在衍生屏上同一位置上，这是理所当然的，只不过人类现有的观测手段还不能判别出这两个本属不同性质的粒子的不同之处，从而误认为相同的粒子会有不同的行为并把它在哲学上归为自然的随机性。这种情况好比有一个外星人降落在地球上时，他们在接近地球人类之前远距离观察地球人时发现：地球人都是完全相同的，因为地球人都有一个头、两支手、两条腿等等；然后，当他们走近地球人时才发现，地球人在身高、相貌上还有更多的区别；当他们和地球人相处一段时间后进一步发现，不同的地球人还存在性别甚至性格上的差别，……，等等。这种认识过程是一个渐进的不断深化的过程。如果穿过窄孔的两个电子还存在除已知的物理特征之外的其他精细特性上的差别，则它们打在衍射屏上的不同位置是完全可以理解的，并且完全符合因果律。事实上，我们对基本粒子的认识过程本来就是渐进的。基本粒子按电荷、质量进行分类在早期的粒子物理学研究中就完成了，但粒子还存在“自旋”这一个特征却是稍晚由泡利发现的。说不准今后什么时候还会发现更加精细的粒子特征区别呢？这就是玻姆的“隐变量”理论，他认为可能存在人们现在还未发现的基本物理参数，使象电子这样的基本粒子还可继续细分。玻姆教授称这种还未发现的参数为“隐变量”或“隐参数”，

他领导着一个物理学家群体长期致力于这种发现“隐变量”的研究工作，据说最近几十年已有重要的发现。

正如物理学家们在微观物理学中所面临的困境一样，博弈论专家也发现混合战略博弈给人们带来了类似的问题。如果我们坚守新古典经济学家们的“理性人”假设，则面临如何解释局中人在无差异的纯战略（恰如物理学中“完全相同”的粒子）之间进行随机选择的问题。一种方案就是采用类似于哥本哈根学派的观点，即假设局中人可以在无差异的纯战略之间“自主地”进行随机选择，好比尼尔斯·玻尔在“饿死的驴”故事中的解决方案，我们假定局中人为了不让其他局中人看出自己的纯战略选择，他以某一概率分布在不同的带来同样的期望支付的纯战略之间进行选择以迷惑其他局中人。推论 3.2.1 中关于混合战略博弈纳什均衡战略中的以正概率进行选择的纯战略带来同样期望支付的结果是显然的，因为如果有两个纯战略具有不同的期望支付，则其他局中人不会认为该局中人会选择期望支付较小的纯战略。而认为该局中人会选择混合战略博弈纳什均衡战略中期望支付最大的那一个纯战略，其他局中人在这种判断下就会选择战胜该局中人选择这一纯战略的相应纯战略，而当该局中人先前也知道这一结果时，他又不会选择开始认为期望支付最大的那个纯战略（他修改他对其他局中人选择不同纯战略的概率分布判断或“信念”，或修改他对其他局中人选取的混合战略的判断），而这种调整又被其他局中人所预知，他们又会继续进行调整，……。这种调整过程会无限止的进行下去，永远也不会达到均衡。例如，在“石头、剪子、布”博弈中，如果局中人 1 选择“石头”的期望支付大于他选择“剪子”和“布”的期望支付，则局中人 2 会认为局中人 1 会选择纯战略“石头”，因而他就选择“布”；局中人 1 知道局中人 2 的这一心理过程，他会改选“剪子”以胜取局中人 2 的“布”，而局中人 2 也预知局中人 1 的这一推理，他又会改选“石头”，……，这一过程是没有止境的。为了达成均衡，局中人 1 在判断局中人 2 的混合战略或概率分布时，只有假设它恰能保证局中人 1 选取“石头”、“剪子”和“布”时都带来同样的期望支付才可能达成均衡；同样，对局中人 2 也是相类似的，他对局中人 1 的概率分布的“信念”（beliefs）必须满足这样的条件：即他无论是选择“石头”、“剪子”还是“布”，这些不同的纯战略都带来同样的期望支付，这样才可能达成均衡。但是，基于理性人假设，局中人只须根据期望支付的大小来作出纯战略选择，当不同的纯战略都带来同样的期望支付时，理性人假设就不能预测局中人到底会选择何种纯战略。摆脱这种困境的思路类似于前面所提到的物理学家们在量子力学中所做的那样。一种方案是改变理性人假设，就象哥本哈根学派所提出的哲学思想那样，用随机性或不可预测性去替换带有机械论色彩的理性人假定，这就好比尼尔斯·玻尔用自然的随机性去取代绝对的因果律——一些博弈论学者认为：我们所以在现实中观测到混合博弈现象（如“石头、剪子、布”游戏），是因为混合战略使一个局中人可以给其他局中人造成不确定性，这样，尽管其他局中人知道他选择某个特定纯战略的概率是多少，但他

们并不能猜透他实际上会选择哪一个纯战略；正是因为他在几个纯战略之间是无差异的，他的行为才难以预测，混合战略博弈均衡才会存在。如果他严格偏好于某个特定的纯战略，他的行为就会被其他局中人准确地猜出，就不会有混合战略博弈均衡出现。这样一种思想的核心是——通过将局中人战略空间的扩大，即战略空间不仅包括纯战略，而且还包括混合战略，纯战略可作为混合战略的一种退化形式；甚至，我们可以将混合战略看成是一个更高层次上的纯战略——局中人在这个扩大了的战略空间中确定性地而不是随机地选择“纯战略”，而其中某些“纯战略”其实是原来较低层次上的混合战略。譬如，如果原来的战略式博弈为 $G=\{S_1,\cdots,S_n;u_1,\cdots,u_n\}$，其中 $S_i\,(i=1,\cdots,n)$ 是纯战略空间，设该博弈的混合战略空间为 $\overline{G}=\{\sum_1,\cdots,\sum_n;v_1,\cdots,v_n\}$，我们在更高层次上将 $\overline{G}$ 看成是纯战略空间为 $\sum_i$、支付函数为 $v_i\,(i=1,\cdots,n)$ 的纯战略博弈，从而 $\overline{G}$ 的“纯战略纳什均衡”就包括了所有 G 的纯战略纳什均衡和混合战略博弈纳什均衡。当然这种扩充还可以进一步进行下去，因为当我们将 $\overline{G}$ 视为纯战略博弈时，它也有进一步的扩充，即 $\overline{\overline{G}}=\{\overline{\sum}_1,\cdots,\overline{\sum}_n,\overline{v}_1,\cdots,\overline{v}_n$，其中 $\overline{\sum}_i$ 为将 $\sum_i$ 中元素视为纯战略时的混合战略空间，即 $\overline{\sum}_i=\{\sigma_i=(\sigma_{i1},\cdots,\sigma_{ik(i)})|\sigma_{ih}$ 是局中人 i 选择 s_{ih} 的概率，$\sum_{h=1}^{k(i)}\sigma_{ih}=1$，$0\le\sigma_{ih}\le1$，$s_{ih}\in\sum_i$，$k(i)$ 为 $\sum_i$ 中元素的个数}，而 $\overline{v}_i$ 是在这个进一步扩充后的博弈中局中人 i 的期望支付，$i=1,\cdots,n$。我们不难证明，这种进一步扩充是没有意义的，因为有 $\overline{\sum}_i=\sum_i$，$\overline{v}_i=v_i\;(i=1,\cdots,n)$，即这种扩充只须一步就行了。这样，我们将 $\overline{G}$ 中的纯战略纳什均衡看作是 G 中的混合战略博弈均衡（包括纯战略博弈均衡）。这可看作是对混合战略博弈的一种解释。

定理 3.4（蒲勇健，2001） 设战略式表述博弈 $G=\{S_1,\cdots,S_n;u_1,\cdots,u_n\}$，其混合战略博弈为 $\overline{G}=\{\sum_1,\cdots,\sum_n,v_1,\cdots,v_n\}$；令 $\overline{\overline{G}}$ 为 $\overline{G}$ 的混合战略博弈，即 $\overline{\overline{G}}=\{\overline{\sum}_1,\cdots,\overline{\sum}_n,\overline{v}_1,\cdots,\overline{v}_n\}$，其中 $\overline{\sum}_i$ 和 $\overline{v}_i$ 的定义是基于如下假设：当我们将 $\overline{G}$ 视为一种纯战略博弈时，$\overline{\sum}_i$ 和 $\overline{v}_i$ 分别是其混合战略空间和期望支付函数。则有 $\overline{\overline{G}}=\overline{G}$。

证明：显然有 $\sum_i\subset\overline{\sum}_i$，$i=1,\cdots,n$，下面证明反过来也有 $\overline{\sum}_i\subset\sum_i$，$i=1,\cdots,n$。

首先，我们给出 $\sum_i$ 中元素的一种刻画，以便于以下的证明。假设纯战略空间 S_i 的集合论表示为 $S_i=\{s_{p_i}\}_{p_i\in P_i}$，$p_i$ 为指标，P_i 为指标集，则 $\forall\sigma_i\in\sum_i$，可有如下表示：

$\sigma_i=\{\sigma(p_i)\}_{p_i\in P_i}$，满足约束条件：

$$\begin{cases}0\le\sigma(p_i)\le1\\ \int_{P_i}\sigma(p_i)=1\end{cases}\tag{3.29}$$

其中 $\sigma(p_i)$ 是局中人 i 在 S_i 中选择 s_{p_i} 的概率。记由（3.29）决定的 $\sigma(p_i)$ 的集合为 H_i，则 $\sum_i$ 的集合论表示为

$\sum_i=\{\sigma_{h_i}\}_{h_i\in H_i}$

$\forall \bar{s}_i \in \overline{\sum}_i$，局中人$i$在$\bar{s}_i$中选择$\sigma_{h_i} \in \sum_i$的概率记为$\bar{\sigma}_{h_i}$，则$\bar{s}_i = \{\bar{\sigma}_{h_i}\}_{h_i \in H_i}$，满足条件：

$$\begin{cases} 0 \le \bar{\sigma}_{h_i} \le 1 \\ \int_{H_i} \bar{\sigma}_{h_i} dh_i = 1 \end{cases} \qquad （3.30）$$

根据定义式（3.6），得

$$\bar{v}_i(\bar{s}_1, \cdots, \bar{s}_n) = \sum_{\sigma \in \Sigma} p(\sigma) v_i(\sigma) \qquad （3.31）$$

其中$\sigma = (\sigma_1, \cdots, \sigma_n)$，$\Sigma = \prod_{i=1}^{n} \Sigma_i$，$p(\sigma)$为混合战略组合$\sigma = (\sigma_1, \cdots, \sigma_n)$出现的概率。在局中人战略选择的独立性假定下，我们有

$$p(\sigma) = \prod_{j=1}^{n} p_j(\sigma_j) \qquad （3.32）$$

其中 $p_j(\sigma_j)$ 为局中人 j 选择混合战略 $\sigma_j \in \sum_j$ 的概率， $j = 1, \cdots, n$ ； 如果 $\sigma_j = \sigma_{h_j}(h_j \in H_j)$

则$p_j(\sigma_j) = \bar{\sigma}_{h_j}$，于是式（3.32）可写为

$$p(\sigma) = \prod_{j=1}^{n} \bar{\sigma}_{h_j} \qquad （3.33）$$

式（3.31）变为

$$\bar{v}_i(\bar{s}) = \sum_{h \in H} \prod_{j=1}^{n} \bar{\sigma}_{h_j} v_i(\sigma_{h_1}, \cdots, \sigma_{h_n}) \qquad （3.34）$$

其中$h = (h_1, \cdots, h_n)$，$H = \prod_{j=1}^{n} H_j$

据式（3.7），得

$$v_i(\sigma_{h_1}, \cdots, \sigma_{h_n}) = \sum_{s \in S} \prod_{j=1}^{n} \sigma_{h_j}(p_j) u_i(s) \qquad （3.35）$$ [15]

其中$s = (s_1, \cdots, s_n)$，$S = \prod_{j=1}^{n} S_j$

式（3.35）也可写成

$$v_i(\sigma_{h_1}, \cdots, \sigma_{h_n}) = \sum_{p \in P} \prod_{j=1}^{n} \sigma_{h_j}(p_j) u_i(p_1, \cdots, p_n) \qquad （3.36）$$

其中，$p = (p_1, \cdots, p_n)$，$P = \prod_{j=1}^{n} P_j$。

将式（3.36）代入式（3.34）：

$$
\begin{aligned}
\bar{v}_i(\bar{s}) &= \sum_{h\in H}\prod_{j=1}^{n}\bar{\sigma}_{h_j}\sum_{p\in P}\prod_{j=1}^{n}\sigma_{h_j}(p_j)u_i(p_1,\cdots,p_n) \\
&= \sum_{p\in P}\left[\sum_{h\in H}\prod_{j=1}^{n}\bar{\sigma}_{h_j}\sigma_{h_j}(p_j)\right]u_i(p_1,\cdots,p_n) \qquad (3.37) \\
&= \sum_{p\in P}A_p u_i(p)
\end{aligned}
$$

其中 $A_p=\sum_{h\in H}\prod_{j=1}^{n}\bar{\sigma}_{h_j}\sigma_{h_j}(p_j)=\prod_{j=1}^{n}\left[\sum_{h_j\in H_j}\bar{\sigma}_{h_j}\sigma_{h_j}(p_j)\right]$

根据定义，$\bar{\sigma}_{h_j}$ 是局中人 j 在 $\bar{s}_j$ 中选择 $\sigma_{h_j}\in\sum_j$ 的概率，而他在 σ_{h_j} 中选择纯战略 $p_j\in S_j$ [16] 的概率为 $\sigma_{h_j}(p_j)$，故局中人 j 在 $\bar{s}_j$ 中选择纯战略 p_j 的概率为 $\sum_{h_j\in H_j}\bar{\sigma}_{h_j}\sigma_{h_j}(p_j)=p_j(p_j)$，其中 $p_j(p_j)$ 为 p_j 出现的概率，于是，将式（3.37）简写为

$$\bar{v}_i(\bar{s})=\sum_{p\in P}\prod_{j=1}^{n}p_j(p_j)u_i(p) \qquad (3.38)$$

或

$$\bar{v}_i(\bar{s})=\sum_{s\in S}p(s)u_i(s) \qquad (3.39)$$

其中 $p(s)=\prod_{j=1}^{n}p_j(p_j)$

局中人 i 在 $\bar{s}_i$ 中选择纯战略 s_{p_i} 的概率为 $\sum_{h_i\in H_i}\bar{\sigma}_{h_i}\sigma_{h_i}(p_i)\geq 0$，又有

$$\int_{P_i}\sum_{h_i\in H_i}\bar{\sigma}_{h_i}\sigma_{h_i}(p_i)dp_i=\sum_{h_i\in H_i}\bar{\sigma}_{h_i}\int_{P_i}\sigma_{h_i}(p_i)dp_i=\sum_{h_i\in H_i}\bar{\sigma}_{h_i}=\int_{H_i}\bar{\sigma}_{h_i}dh_i=1$$

故 $\bar{s}_i\in\sum_i$，于是有

$\overline{\sum}_i\subset\sum_i$， $i=1,\cdots,n$

再结合式（3.39），据定义式（3.6）可知 $\bar{v}_i(\bar{s})=v_i(\bar{s})$，于是定理证毕！

基于定理 3.4，我们似乎可以对混合战略博弈给出一种解释。尽管从较低层次的博弈 G 来看，混合战略博弈纳什均衡战略表现为局中人在一些无差异的纯战略之间进行随机性战略选择，但从较高层次的博弈 $\overline{G}$ 来看，这种选择却是确定性的（我们在这里将 $\overline{G}$ 看成是一种纯战略博弈）。因为，在 G 中选择一个混合战略 $\sigma_i\in\sum_i$ 对应于在 $\overline{G}$ 中选择一个纯战略 $\bar{s}_i\in\overline{\sum}_i$。当 $\bar{s}_i\in\overline{\sum}_i$ 是一个 $\overline{G}$ 中的混合战略时，根据定理 3.4，我们知道 $\bar{s}_i$ 又是 G 中的一个混合战略，即定理 3.4 的意思是说 G 的混合战略空间 $\sum_i$ 在混合战略运算下是封闭的（用数学语言来说就是：$\sum_i$ 是 S_i 的凸包，而一个集合 $\sum_i$ 的凸包的凸包就是集合 $\sum_i$ 本身），所以，$\bar{s}_i$ 又是 $\overline{G}$ 中的纯战略，也就是说，$\overline{G}$ 中的混合战略都是 $\overline{G}$ 中的纯战略；对于博弈 $\overline{G}$ 来说，可以不存在混合战略与纯战略之间的区分。所以，可以认为 G 中的混合战略博弈就是 $\overline{G}$ 中的纯战略博弈，局中人在 $\overline{G}$ 中的战略选择完全是基于理性人假设的确定性行为。如果我们考虑 $\overline{G}$ 中的（强）纳什均衡，则局中人的均衡战略选择完全是

基于不同战略带来的（期望）支付之间的差别来进行的。这样，随机性问题就完全消除了，我们可以又回到博弈论所基于的理性人假设上去，博弈论的基础不因混合战略带来的哲学问题而动摇。但是，这种解释仍面临两个方面的质疑。首先，我们难以安然地接受一种看来是互相矛盾的现实——在较高层次的博弈$\overline{G}$中，局中人确定性而不是随机性地进行战略选择，但在较低层次的博弈G中，局中人同时又是随机性而不是确定性地进行战略选择，有什么理由使我们接受$\overline{G}$中的确定性行为而忽视G中的随机性行为呢？我们凭什么说因为存在$\overline{G}$中的确定性行为而认为局中人的战略选择就是确定性的呢？为什么不能因为G中的随机性行为而认为局中人的战略选择是随机性的呢？其次，基于定理 3.4，我们从$\overline{G}$的角度可以将G中的混合战略视为一种纯战略，这种理解的核心解说是：局中人在博弈中随机地“虚张声势”本身是一种战略，一个局中人选择混合战略的目的是给其他局中人造成不确定性，这样，尽管其他局中人知道他选择某个特定纯战略的概率是多少，但他们并不能猜透他实际上会选择哪个纯战略。事实上，正是因为他在几个战略之间是无差异的，他的行为才难以预测，混合战略均衡才会存在。如果他严格偏好于某个特定的纯战略，他的行为就会被其他局中人准确地猜透，就不会有混合战略均衡出现（张维迎，1996，第 108 页）。这样，政府可以随机地稽核纳税人，一些商店也随机地提供折扣等。但是，这种解释面临一种深刻的质疑是：一般地局中人深思熟虑地去随机化是为了影响别的局中人的行动。例如，考虑孩子们玩的猜谜游戏（见例 3.13 和例 3.15），其中局中人选择是出硬币的正面还是反面；这个博弈被经典地用来说明混合战略均衡这一概念的动机，但是随机性是关于博弈中局中人深思熟虑战略的奇怪描写。局中人的行动是对别人选择的猜测的反应；“猜”是一种极精细而非随机的心理活动（见 MJ Osborne 与 A Rubinstein，1994，第 3 章第 2 节）。这就是说，在事实上而不是理论上来看，猜测是基于深刻分析已有信息的确定性行为而不是随机性的。

类似于爱因斯坦——玻姆的决定论哲学思想，另外一些博弈论专家试图从“隐变量”角度思考对混合战略博弈的解释。在现实生活中，影响局中人战略选择的因素是相当多的，有时多到连局中人也难以描述到底是哪些因素以什么方式影响他自己的行为，当然，这时对于建模者来说，就更加难以在模型中将这些众多因素包括进来。概率论的一种基本思想就是，当影响某一变量的因素足够多时，我们就可以将这一变量视为“随机变量”，即假设这一变量的变动无法准确预测，最多也只能预测其变动的概率分布。譬如，士兵射击的准确度受到许多因素的影响——士兵射击姿势，风向风力，扣板机时的用力度，……，等等，而其中任何一种因素如士兵射击姿势，风向风力等又受到其它更多因素的影响，我们不可能用力学方程将所有这些因素的具体影响都一一写出来，只有假定射击准确度（如环数）是一个随机变量。但是，在一个特定的环境下，尽管影响射击准确度的因素很多，但这些因素的作用方式却是相当稳定的，总体上表现为射击准确度这个随机变量的分布却是确定的。概率的概念只有在多次重复的活动中才有意义，如掷硬

币实验，当掷出一万次硬币中大约有 5000 次掷出正面，我们就说正面出现的概率为 0.5。如果我们基于这种思路去寻找对混合战略博弈的解释，我们就必须将所讨论的“静态博弈”扩展到本书后面章节中将要介绍的（重复性）动态博弈中去。假设一个（一次性）博弈重复许多次（如小孩们进行多次猜谜），在每一次博弈中，影响局中人的其它众多没有被写入模型的因素对局中人的影响有所不同，因而每一次博弈尽管是纯战略博弈，其结果也是纯战略均衡，但不同次博弈的纯战略均衡却不一样，表现为在不同次博弈中局中人选择的纯战略是不同的。当这种重复性博弈重复足够多次数后，我们会发现局中人选择不同纯战略的频数（即选择纯战略的次数与博弈重复次数之比）可能是“稳定的”——这种频数不随博弈次数的增加而显著变动，恰如在掷硬币实验中正面出现的频数那样，我们可将这个频数（或这个频数在博弈重复次数趋于无限时的极限）视为一种 “混合战略”中局中人选择该纯战略的概率，并将其作为预测局中人在每一次（静态）博弈中选择该纯战略的概率；这样，我们就获得对混合战略博弈的另一种解释，它基于类似于物理学中的“隐变量”理论——玻姆用它来对基本粒子的“随机性”行为作出本质上基于决定论的解释——来对局中人的“随机性”战略选择作出解释。显然，这种解释本质上是一种工具主义，因为它对混合战略博弈模型的解释是通过基于确定性假设但工具性地用混合战略博弈刻画局中人在复杂环境中的动态行为。这种解释的出发点是刻画复杂环境中的众多因素对局中人的影响机制是需付出高成本的，因而通过混合战略博弈这一人为制造的（也许不是真实的）模型去刻画博弈显得更加“经济”一些。这种解释的基础是：混合战略博弈均衡必须作为真实的重复博弈过程的“稳定状态”（Steady–State），即随着博弈重复次数的增加，局中人选择任一给定纯战略的频数趋于一个确定的极限数。同时，这种解释还意味着在每一阶段博弈（即构成重复性博弈的每一次一次性“静态”博弈）中局中人选择的纯战略组合将成为在当前环境下的纯战略纳什均衡，而正是局中人在不同次阶段博弈中的环境变化导致局中人在不同次阶段博弈中选择不同的纯战略，将这种纯战略均衡的变动极限转变为“混合战略均衡”正是这种解释的核心思想（也许我们可从这种纯战略均衡的变动或从混合战略均衡中可以反求局中人在不同次阶段博弈中所处环境的变动，但结果可能不是唯一的）。通过改变局中人在每一次阶段博弈中的支付函数，我们可以保证阶段博弈中任何纯战略组合都是纳什均衡，这种改变或许恰好反映了该阶段博弈中相对于以前阶段已经发生了的环境变化，但事实上没有人能证明事实正是如此。

基于这种理论，我们能将一个混合战略纳什均衡解释为一个随机稳定状态。局中人拥有过去行动被采用频数的信息（如“局中人作为局中人 1 的角色在 80%的阶段博弈中选择战略 s_{11}，在 20%的阶段博弈中选择战略 s_{12}”），每个局中人使用这些频数信息去形成他的关于别的局中人未来战略选择的信念，因此可以系统地决定他自己的战略选择。在均衡中这些频数随时间保持不变并且在这样的意义下是稳定的：给定稳定状态信念，

由局中人用正概率选择的任何战略都是最优的。

混合战略均衡预示着博弈的结果是随机的（尽管它是基于确定性决策的结果，随机性仅仅反映了局中人面临的环境因素变化的随机性），所以对博弈的单个行动来说它的预测比纯战略均衡的预测精确性要差。但是，理论的作用是用来解释规则的；混合战略均衡的概念抓住了随机性规则。

基于这种解释，我们可以如此理解混合战略博弈。在选择他的战略前，一个局中人可能收到他的战略或许要依赖的随机个人信息，但从别的局中人角度看，这些信息并不重要。局中人可能不会有意识地选择利用这种存在于他的战略与他得到的个人信息间的联系，这种联系也许仅是使他的行动被其他的局中人或外部观察者看来似乎是“随机的”。在把局中人的行为模型化为随机的过程中，混合战略纳什均衡抓住了有关局中人视作不相关因素方面的行为依存关系。换言之，局中人可能知道决定他对手行为的外部因素，但是又可能发现要确定这种关系是不可能的或是成本相当高的（出于同样理由，我们将掷硬币结果模型化为随机的，而不将它描述成以下因素相互作用的结果：它的初始位置和速度、风速及其它因素等）。简言之，从这个方式来看混合战略纳什均衡是对系统稳定状态的描述，这个系统反映了博弈初始描述所缺失的元素（MJ Osborne 和 M Rubinstein，第 3 章第 3 节）。

我们通过一个例子来说明上述设想。在 3.2.1 小节所介绍的“性别战”博弈中，如我们已知，此博弈有一个混合战略纳什均衡$\left\{\left(\frac{2}{3},\frac{1}{3}\right),\left(\frac{1}{3},\frac{2}{3}\right)\right\}$。现在假设每个局中人有三种由他所不知的因素确定的“心情”。每个局中人在每个心情里呆$\frac{1}{3}$时间，并与别的局中人心情无关；他（她）的心情不影响他（她）的支付。假设：无论在心情 1 还是在心情 2，局中人“男”选择“足球”，当在心情 3 时选择“芭蕾”；在心情 1 局中人“女”选择“足球”，在心情 2 或心情 3 下选择“芭蕾”。将此情形视为每个局中人三种类型对应于他（她）的可能心情的博弈，这种博弈假设每个局中人都不知道对方所处于的“心情”，只知道对方可能处于各种心情的概率分布，也知道对方知道自己处于各种“心情”的概率分布，……，即局中人“心情”的概率分布是“共同知识”。这种假设下的静态博弈被称为“不完全信息静态”博弈或贝叶斯博弈（Bayes game）。此种设定确定了一个恰与原来的“性别战”博弈的混合战略纳什均衡相对应的纯战略纳什均衡。请注意对这个混合战略均衡的解释不依赖于每个局中人具有三个相同可能性的且独立的心情，我们仅需要局中人个人信息满足它们能产生恰当的随机变量。

但是，对于这种解释至少存在一种批评意见。局中人深思熟虑的战略选择依赖于不影响他（她）支付的因素是很难令人接受的。因为人们通常为他们的选择给出理由；在任何特定情形下，希望应用混合战略均衡概念的建模者都应该指出支付不相关的理由及

解释局中人个人信息与他（她）的选择间所需要的依存关系。对于这种批评，一种可能的回答是：混合战略均衡中每个人对他（她）的均衡战略支集（数学家们称使某函数不为零的自变量集合为该函数的支集，这里指混合战略中以正概率选择的那些纯战略所构成的集合）上的所有战略具有相同偏好，所以说选择的行动依赖于被建模者认为“不相关的”因素不是没有道理的。当被问到他们为什么从每个元素都一样具有吸引力的集合中选定一个战略时，人们经常给出诸如：“我不知道——我只是觉得喜欢它”之类的答案（但这种回答显然与理性人假设相冲突，它更多地类似于随机性解释）。

无论怎样，这种解释的关键假设是局中人选择的战略可能依赖于模型之外的因素，即将混合战略解释为一个更大博弈中的纯战略。

目前，对混合战略均衡的最具影响力的解释是由 Harsanyi（1973）提出的。Harsanyi 理论在许多方面综合了前面提到的“隐变量”或外生因素解释的多种观点，并基于信息不完全条件下表达了这些因素如何导致局中人采用随机方式选择纯战略的机理。

在 Harsanyi 的理解里，混合战略均衡产生自局中人的类型的不确定性，这种不确定性当然又产生自多种原因，如上文所述的模型未能刻画的其它许多因素。当局中人的类型如支付函数存在不确定性时，局中人在纯战略选择上尽管是深思熟虑的并且给定每一种确定的场合都是基于理性人假设下的理性选择（严格最优战略选择），但由于他自身在类型上的多种可能性以及他对其他局中人类型的不确定性认识，使得他的选择在确定性的支付函数表述下的模型中表现为好象是随机的。假定每个局中人类型的分布函数是共同知识，但任何一个局中人都不知道其他局中人的具体类型实现，因而其他局中人的类型对他来说是一种贝叶斯（Bayes）统计意义上的“随机变量”，他为了其期望支付最大化，在战略选择上就表现为随机的；同时，由于他自身的类型也是不确定的，尽管他自己总是知道自己类型的实现，但其他人（包括建模者）不知道他的类型实现，因而使他的战略选择就更加表现出随机的性质。

令 $G=\{S_1,\cdots S_n;u_1,\cdots,u_n\}$ 为一有限博弈，$\bar{\varepsilon}=\{\varepsilon_i(s)|_{i=1,\cdots,n},s\in S\}$ 为区间[–1, 1]中的一族随机变量，这里 $\varepsilon_i=\{\varepsilon_i(s)\}_{s\in S}$ 有一个连续可微的密度函数和一个绝对连续的分布函数，并且随机向量 $(\varepsilon_i)i=1,\cdots,n$ 是独立的。考虑一组不确定的博弈，其中每个局中人 i 在结果 s 的支付服从小随机变化 $\varepsilon_i(s)$，每个局中人 i 知道 ε_i 的实现值为 $\{\varepsilon_i(s)\}_{s\in S}$，但不知道别的局中人随机变量的实现情况。这样，可考虑贝叶斯博弈 $G(\bar{\varepsilon})$，其中自然状态集合为 $\bar{\varepsilon}$ 所有可能实现值的集合，每个局中人的（共同）先验概率（即局中人事先关于有关随机变量分布函数的信念）是由 $\bar{\varepsilon}$ 确定的概率分布，局中人 i 仅知道 $\{\varepsilon_i(s)\}_{s\in S}$ 的实现情况及局中人在结果为 s 和状态为 $\bar{\varepsilon}$ 情况下他的支付是 $u_i(s)+\varepsilon_i(s)$。

Harsanyi 的主要结论（1973，定理 2 和 7）是：对几乎[17]任一博弈 G 和任一满足上述条件的随机变量族 $\bar{\varepsilon}^*$，几乎任一 G 的混合战略纳什均衡都是与某个极限相联系的混合战略组合，而该极限为贝叶斯博弈 $G(\gamma\bar{\varepsilon}^*)$ 的纯战略均衡序列，随不确定性的大小，γ 逐

渐消失的极限，当然$G(\gamma\bar{\varepsilon}^*)$应满足对于$G(\gamma\bar{\varepsilon}^*)$中的每个博弈而言由每个类型的局中人所选择的行动都是严格最优的。进一步，任何这种收敛序列的极限都与G的一个混合战略均衡相联系（Harsanyi，1973，定理 5）。也就是，当支付中的随机变化小的时候，博弈G的几乎任一混合战略均衡都接近于相联系的贝叶斯博弈的纯战略均衡，反之亦然。数学家们称具有这种性质的G的混合战略均衡在$\bar{\varepsilon}^*$下是“可逼近的”（approachable）。由于与这些结论相关的数学复杂性，我们在这里不给出证明。

3.2.13 混合战略空间的重复剔除的占优战略均衡

3.2.12 小节中的定理 3.4 告诉我们，混合战略可被视为一个更高层次上的博弈中的纯战略，而对于纯战略博弈，我们已有关于占优均衡和重复剔除的占优战略均衡的许多概念和结果。这里，一个十分自然的想法是：能否在这个更高层次的博弈中推广和应用这些概念和结果？答案是肯定的。首先，在混合战略情况下，关于剔除劣战略和重复剔除劣战略的概念及结论依然成立，即任何局中人都不会选择任何的严格劣战略，重复剔除严格劣战略的顺序与剔除严格劣战略后剩下的博弈无关，当博弈是重复剔除占优可解的时，剔除劣战略的顺序与重复剔除的占优战略均衡无关等等。

例 3.22 给出了一个在混合战略空间进行求解重复剔除劣战略的占优战略均衡的例子。

例 3.22

表 3.22 给出的博弈中，局中人 1 有三种战略，局中人 2 有两种战略。

表 3.22　混合战略空间中的重复剔除劣战略

		2	
		L	R
1	U	3, 1	0, 2
	M	0, 2	3, 3
	D	1, 3	1, 1

从 2 的角度看，L 和 R 都不是他的严格劣战略，因为当 1 选 U、M 时，他选 R 的支付大于选 L 的支付，而 1 选 D 时他选 L 的支付大于选 R 的支付。再比较局中人 1 的三种战略，同样可知 U、M、D 三种战略两两之间没有哪一个绝对比另一个好或绝对比另一个差，因此，倘若我们只考虑纯战略博弈，局中人 1 也没有任何的严格劣战略。但是，如果我们允许局中人 1 选取混合战略，则情况就不一样了。设局中人 1 选取混合战略$\left(\frac{1}{2},\frac{1}{2},0\right)$，则与这个混合战略相比，D 一定是 1 的严格劣战略。事实上，当 2 选取纯战略 L 时，1 的混合战略支付为：

$\frac{1}{2}\cdot 3+\frac{1}{2}\cdot 0+0\cdot 1=\frac{3}{2}$

当 2 选取纯战略 R 时，1 的混合战略支付为：

$\frac{1}{2}\cdot 0+\frac{1}{2}\cdot 3+0\cdot 1=\frac{3}{2}$

即使 2 也选取某个混合战略$(\gamma,1-\gamma)$，1 选取上述混合战略的支付仍为：

$\frac{1}{2}\cdot\gamma\cdot 3+\frac{1}{2}\cdot(1-\gamma)\cdot 0+\frac{1}{2}\cdot\gamma\cdot 0+\frac{1}{2}\cdot(1-\gamma)\cdot 3=\frac{3}{2}$

即无论 2 选取什么战略（纯战略和混合战略），1 选这个混合战略获得的期望支付总大于他选纯战略 D 时获得的确定性支付，如果我们假设 1 是风险中性的[18]，则纯战略 D 是严格劣战略。

于是，我们将 D 剔除，得到表 3.23 表示的博弈。此时很容易看出，L 是 2 相对于 R 的严格劣战略，可将它从 2 的战略空间中剔除。最后，在剩下的博弈中，1 当然选 M，2 选 R，得到重复剔除的占优战略均衡：(M, R)。

表 3.23　剔除 D 后的博弈

		2	
		L	R
1	U	3, 1	0, 2
	M	0, 2	3, 3

显然，定理 3.1 及其推论均适合于混合战略博弈的情形，因为当我们将定理 3.4 中的博弈$\overline{G}$取代定理 3.1 中的博弈G后，定理 3.1 的证明过程不会有实质性的变化，同样，推论 3.1.1 也可推广到$\overline{G}$的情形。

3.3 几种均衡概念的涵盖与交叠关系

我们已经介绍的在博弈论发展史上相继提出的几种不同的均衡概念，即最大最小均衡、占优战略均衡、重复剔除的占优战略均衡和纳什均衡。现在关心的问题是：这几种不同的均衡概念之间是什么样的一种关系？显然，占优战略均衡是重复剔除的占优战略均衡的一种特例，不太显然的是，重复剔除的占优战略均衡并不是纳什均衡的一种特例，某些教科书[19]将重复剔除的占优战略均衡说成是纳什均衡的一种特例是不正确的。事实上，存在某些重复剔除的占优战略均衡，但它们并不是纳什均衡。为了说明这一点，我们在这里给出一个反例。

例 3.23　设战略式博弈$G=\{S_1,S_2;u_1,u_2\}$，其中$S_i(i=1,2)$的定义如下：

$S_1=S_2=\left\{0,\cdots,\frac{1}{n},\cdots,\frac{1}{2},1\right\}$，其中$n$为自然数；支付函数定义如下：

$u_1(0,0)=0$；

$u_1\left(0,\frac{1}{n}\right)=u_2\left(\frac{1}{n},0\right)=A$，A 是充分大的正数

$u_1\left(\frac{1}{n},0\right)=u_2\left(0,\frac{1}{n}\right)=\varepsilon+1+\cdots+\frac{1}{2^n}$；这里 n 为不等于 1 的自然数；

$u_1\left(1,\frac{1}{n}\right)=u_2\left(\frac{1}{n},1\right)=u_1(1,0)=u_2(0,1)=\varepsilon$；

$u_1\left(\frac{1}{2},\frac{1}{n}\right)=u_2\left(\frac{1}{n},\frac{1}{2}\right)=u_1\left(\frac{1}{2},0\right)=u_2\left(0,\frac{1}{2}\right)=\varepsilon+1+\frac{1}{2}+\frac{1}{2^2}$；

$u_1\left(\frac{1}{3},\left.\frac{1}{n}\right|_{n\neq1}\right)=u_2\left(\left.\frac{1}{n}\right|_{n\neq1},\frac{1}{3}\right)=u_1\left(\frac{1}{3},0\right)=u_2\left(0,\frac{1}{3}\right)=\varepsilon+1+\frac{1}{2}+\frac{1}{2^2}+\frac{1}{2^3}(<\varepsilon+2)$

$u_1\left(\frac{1}{3},1\right)=u_2\left(1,\frac{1}{3}\right)=\varepsilon+\frac{1}{2^3}$

……

$$u_1\left(\left.\frac{1}{k}\right|_{k\geq3},\left.\frac{1}{n}\right|_{n\geq k-1}\right)=u_2\left(\left.\frac{1}{n}\right|_{n\geq k-1},\left.\frac{1}{k}\right|_{k\geq3}\right)=u_1\left(\left.\frac{1}{k}\right|_{k\geq3,0}\right)=u_2\left(0,\left.\frac{1}{k}\right|_{k\geq3}\right)$$
$$=\varepsilon+1+\frac{1}{2}+\frac{1}{2^2}+\cdots+\frac{1}{2^k}$$

$$u_1\left(\left.\frac{1}{k}\right|_{k\geq3},\left.\frac{1}{n}\right|_{1\leq n\leq k-1}\right)=u_2\left(\left.\frac{1}{n}\right|_{1\leq n\leq k-1},\left.\frac{1}{k}\right|_{k\geq3}\right)=\varepsilon+\frac{1}{2^k}$$

……

其中 n,k 除特别声明外为全体自然数，ε 为任意正数。

下面，我们证明该博弈是重复剔除占优可解的，且重复剔除的占优战略均衡为{0, 0}。以下运用数学归纳法证明 $S_i(i=1,2)$ 中的所有战略元素{$\frac{1}{n}$，n 为自然数}都被剔除，只剩下唯一的一个战略元素{0}。

事实上，当 n 为自然数时，我们有

$$u_1\left(1,\frac{1}{n}\right)=u_1(1,0)=\varepsilon<u_1\left(\frac{1}{2},\frac{1}{n}\right)=u_1\left(\frac{1}{2},0\right)=\varepsilon+1+\frac{1}{2}+\frac{1}{2^2}$$

故局中人 1 的战略“1”被剔除；

因 $u_1\left(0,\frac{1}{n}\right)=A$，$n$ 为自然数，当 A 充分大时，{0}不是劣战略，因为对任何自然数 k,n 都有 $u_1\left(\frac{1}{k},\frac{1}{n}\right)$<A（只要 A 充分大），对任意的自然数 $k\geq3$，有

$$u_1\left(\frac{1}{2},1\right)=\varepsilon+1+\frac{1}{2}+\frac{1}{2^2}>u_1\left(\frac{1}{k},1\right)=\varepsilon+\frac{1}{2^k},$$

所以“$\frac{1}{2}$”不是劣战略。

对任意的自然数$k \ge 3$

当任意自然数$k' > k$，有

$$u_1\left(\frac{1}{k},1\right)=\varepsilon+\frac{1}{2^k}>u_1\left(\frac{1}{k'},1\right)=\varepsilon+\frac{1}{2^{k'}}$$

当任意满足不等式$2 \le k' < k$，有

$$u_1\left(\frac{1}{k},\frac{1}{k}\right)=\varepsilon+1+\frac{1}{2}+\cdots+\frac{1}{2^k}>u_1\left(\frac{1}{k'},\frac{1}{k}\right)=\varepsilon+1+\frac{1}{2}+\cdots+\frac{1}{2^{k'}}$$

当$k'=1$，则

$$u_1\left(\frac{1}{k},\frac{1}{k}\right)=\varepsilon+1+\cdots+\frac{1}{2^k}>u_1\left(1,\frac{1}{k}\right)=\varepsilon$$

所以，“$\frac{1}{k}$”不是局中人 1 的劣战略。

故在第一轮剔除劣战略中，只剔除了局中人 1 的战略空间S_1中的劣战略“1”。

在第二轮剔除劣战略中，我们对局中人 2 的战略空间S_2中的劣战略进行剔除。

因

$$u_2\left(\frac{1}{n},1\right)=u_2(0,1)=\varepsilon<u_2\left(\frac{1}{n},\frac{1}{2}\right)=u_2\left(0,\frac{1}{2}\right)=\varepsilon+1+\frac{1}{2}+\frac{1}{2^2}$$

故局中人 2 的劣战略“1”被剔除；

因有

$$u_2\left(\left.\frac{1}{n}\right|_{n\ge 2},\frac{1}{2}\right)=\varepsilon+1+\frac{1}{2}+\frac{1}{2^2}<u_2\left(\left.\frac{1}{n}\right|_{n\ge 2},\frac{1}{3}\right)=\varepsilon+1+\frac{1}{2}+\frac{1}{2^2}+\frac{1}{2^3}$$

$$u_2\left(0,\frac{1}{2}\right)=\varepsilon+\frac{1}{2}+\frac{1}{2^2}<u_2\left(0,\frac{1}{3}\right)=\varepsilon+1+\frac{1}{2}+\frac{1}{2^2}+\frac{1}{2^3}$$

故局中人 2 的另一劣战略“$\frac{1}{2}$”也被剔除；

显然，{0}不是劣战略。

对任意自然数$k \ge 4$，我们有

$$u_2\left(\frac{1}{2},\frac{1}{3}\right)=\varepsilon+1+\frac{1}{2}+\frac{1}{2^2}+\frac{1}{2^3}>u_2\left(\frac{1}{2},\frac{1}{k}\right)=\varepsilon+\frac{1}{2^k}$$

所以，局中人 2 的战略“$\frac{1}{3}$”不是劣战略；

又对任意的自然数$k \ge 4$，

设k'为自然数

当$k' \ge k$，有

$$u_2\left(\frac{1}{2},\frac{1}{k}\right)=\varepsilon+\frac{1}{2^k}>u_2\left(\frac{1}{2},\frac{1}{k'}\right)=\varepsilon+\frac{1}{2^k},$$

当$k'=1$，有

$$u_2\left(\frac{1}{k},\frac{1}{k}\right)=\varepsilon+1+\cdots+\frac{1}{2^k}>u_2\left(\frac{1}{k},1\right)=\varepsilon$$

当 $k'=2$，有

$$u_2\left(\frac{1}{k},\frac{1}{k}\right)=\varepsilon+1+\cdots+\frac{1}{2^k}>u_2\left(\frac{1}{k},\frac{1}{2}\right)=\varepsilon+1+\frac{1}{2}+\frac{1}{2^2}$$

当 $3\le k'<k$，有

$$u_2\left(\frac{1}{k},\frac{1}{k}\right)>\varepsilon+1+\cdots+\frac{1}{2^k}>u_2\left(\frac{1}{k},\frac{1}{k'}\right)=\varepsilon+1+\cdots+\frac{1}{2^{k'}}$$

故局中人 2 的战略“$\frac{1}{k}$”不是劣战略（$k\ge 3$）。

显然，按这种方法可以将 $S_i(i=1,2)$ 中的战略{$\frac{1}{n}$，n 为自然数}全部剔除，下面运用数学归纳法证明这一点。我们在下面运用定理 3.1 证明过程中所使用的剔除劣战略方法进行劣战略剔除。

首先，第一轮剔除在 $S_i(i=1,2)$ 中将劣战略“1”剔除；根据前面的分析及局中人之间的对称性知 $\left\{\frac{1}{n},n\ne 1\right\}$ 和{0}不是劣战略；

得　$S_i(1)=S_i\quad\{1\}=\{0,\cdots,\frac{1}{n},\cdots,\frac{1}{2},n$ 为不等于 1 的自然数},　i=1,2

因有

$$u_1\left(\frac{1}{2},\frac{1}{n}\Big|_{\ge 2}\right)=u_1\left(\frac{1}{2},0\right)=u_2\left(\frac{1}{n}\Big|_{n\ge 2},\frac{1}{2}\right)=u_2\left(0,\frac{1}{2}\right)=\varepsilon+1+\frac{1}{2}+\frac{1}{2^2}$$

$$<u_1\left(\frac{1}{3},\frac{1}{n}\Big|_{n\ge 2}\right)=u_2\left(\frac{1}{n}\Big|_{n\ge 2},\frac{1}{3}\right)=u_1\left(\frac{1}{3},0\right)=u_2\left(0,\frac{1}{3}\right)=\varepsilon+1+\frac{1}{2}+\frac{1}{2^2}+\frac{1}{2^3}$$

故 $S_i(1)(i=1,2)$ 中的战略“$\frac{1}{2}$”是劣战略，将被剔除；

对任意自然数 $k>3$，有

$$u_1\left(\frac{1}{k},\frac{1}{2}\right)=\varepsilon+\frac{1}{2^k}<u_1\left(\frac{1}{3},\frac{1}{2}\right)=\varepsilon+1+\frac{1}{2}+\frac{1}{2^2}+\frac{1}{2^3}$$

$$u_2\left(\frac{1}{2},\frac{1}{k}\right)=\varepsilon+\frac{1}{2^k}<u_2\left(\frac{1}{2},\frac{1}{3}\right)=\varepsilon+1+\frac{1}{2}+\frac{1}{2^2}+\frac{1}{2^3}$$

故 $S_i(1)(i=1,2)$ 中的战略“$\frac{1}{3}$”不是劣战略；

对任意自然数 $k>3$，

设 k' 是任意自然数，

当 $k'>k$，则

$$u_1\left(\frac{1}{k'},\frac{1}{2}\right)=\varepsilon+\frac{1}{2^{k'}}<u_1\left(\frac{1}{k},\frac{1}{2}\right)=\varepsilon+\frac{1}{2^k}$$

$$u_2\left(\frac{1}{2},\frac{1}{k'}\right)=\varepsilon+\frac{1}{2^{k'}}<u_2\left(\frac{1}{2},\frac{1}{k}\right)=\varepsilon+\frac{1}{2^{k}}$$

当$2\le k'<k$

若$k'=2$，有

$$u_1\left(\frac{1}{k'},\frac{1}{k}\right)=\varepsilon+1+\cdots\frac{1}{2^{k}}>u_1\left(\frac{1}{2},\frac{1}{k}\right)=\varepsilon+1+\frac{1}{2}+\frac{1}{2^{2}}$$

$$u_2\left(\frac{1}{k},\frac{1}{k'}\right)=\varepsilon+1+\cdots+\frac{1}{2^{k}}>u_2\left(\frac{1}{k},\frac{1}{2}\right)=\varepsilon+1+\frac{1}{2}+\frac{1}{2^{2}}$$

若$k'\ge 3$，有

$$u_1\left(\frac{1}{k},\frac{1}{k}\right)=\varepsilon+1+\cdots+\frac{1}{2^{k}}>u_1\left(\frac{1}{k'},\frac{1}{k}\right)=\varepsilon+1+\cdots+\frac{1}{2^{k'}}$$

$$u_2\left(\frac{1}{k},\frac{1}{k}\right)=\varepsilon+1+\cdots+\frac{1}{2^{k}}>u_2\left(\frac{1}{k},\frac{1}{k'}\right)=\varepsilon+1+\cdots+\frac{1}{2^{k'}}$$

故$S_i(1)(i=1,2)$中战略“$\frac{1}{k}$”不是劣战略；

显然，$S_i(2)(i=1,2)$中战略“0”也不是劣战略。

得

$$S_i(2)=S_i(1)\Big\backslash\left\{\frac{1}{2}\right\}$$

下面作归纳法假设，

设$k\ge 2$时，有

$$S_i(k)=S_i(k-1)\Big\backslash\left\{\frac{1}{k}\right\}$$

往证

$$S_i(k+1)=S_i(k)\Big\backslash\left\{\frac{1}{k+1}\right\},i=1,2$$

因

$$\begin{aligned}u_1\left(\frac{1}{k+1},\left.\frac{1}{n}\right|_{n\ge k+1}\right)&=u_2\left(\left.\frac{1}{n}\right|_{n\ge k+1},\frac{1}{k+1}\right)=\varepsilon+1+\cdots+\frac{1}{2^{k+1}}\\&<u_1\left(\frac{1}{k+2},\left.\frac{1}{n}\right|_{n\ge k+1}\right)=u_2\left(\left.\frac{1}{n}\right|_{n\ge k+1},\frac{1}{k+2}\right)\\&=\varepsilon+1+\cdots+\frac{1}{2^{k+2}}\end{aligned}$$

和

$$\begin{aligned}u_1\left(\frac{1}{k+1},0\right)&=u_2\left(0,\frac{1}{k+1}\right)=\varepsilon+1+\cdots+\frac{1}{2^{k+1}}\\&<u_1\left(\frac{1}{k+2},0\right)=u_2\left(0,\frac{1}{k+2}\right)=\varepsilon+1+\cdots+\frac{1}{2^{k+2}}\end{aligned}$$

故 $S_i(k)$ 中的战略“$\frac{1}{k+1}$”是劣战略，$i=1,2$

对任意自然数 $h>0$

$$u_1\left(\frac{1}{k+2+h},\frac{1}{k+1}\right)=\varepsilon+\frac{1}{2^{k+2+h}}=u_2\left(\frac{1}{k+1},\frac{1}{k+2+h}\right)$$

$$<u_1\left(\frac{1}{k+2},\frac{1}{k+1}\right)=u_2\left(\frac{1}{k+1},\frac{1}{k+2}\right)=\varepsilon+1+\cdots+\frac{1}{2^{k+2}}$$

故 $S_i(k)$ 中的战略“$\frac{1}{k+2}$”是劣战略，$i=1,2$

对任意自然数 k'，

当 $k'>k+2+h$，有

$$u_1\left(\frac{1}{k+2+h},\frac{1}{k+1}\right)=u_2\left(\frac{1}{k+1},\frac{1}{k+2+h}\right)=\varepsilon+\frac{1}{2^{k+2+h}}$$

$$>u_1\left(\frac{1}{k'},\frac{1}{k+1}\right)=u_2\left(\frac{1}{k+1},\frac{1}{k'}\right)=\varepsilon+\frac{1}{2^{k'}}$$

当 $k+1\le k'<k+2+h$，有

$$u_1\left(\frac{1}{k+2+h},\frac{1}{k+2+h}\right)=\varepsilon+1+\cdots+\frac{1}{2^{k+2+h}}$$

$$=u_2\left(\frac{1}{k+2+h},\frac{1}{k+2+h}\right)>u_1\left(\frac{1}{k'},\frac{1}{k+2+h}\right)$$

$$=u_2\left(\frac{1}{k+2+h},\frac{1}{k'}\right)=\varepsilon+1+\cdots+\frac{1}{2^{k'}}$$

故 $S_i(k)$ 中战略“$\frac{1}{k+2+h}$”不是劣战略，$i=1,2$；显然，$S_i(k)$ 中战略“0”也不是劣战略；

得

$$S_i(k+1)=S_i(k)\setminus\left\{\frac{1}{k+1}\right\},i=1,2$$

于是有

$$\bigcap_{k=1}S_i(k)=\{0\},\ i=1,2$$

所以，重复剔除的占优战略均衡为{0, 0}，但{0, 0}显然不是纳什均衡，因为有 $u_1(0,0)=0<u_1\left(\frac{1}{n},0\right)$，$n$ 为自然数。

证毕！

尽管一般来说重复剔除的占优战略均衡并不一定是纳什均衡，但对于局中人的战略空间中只有有限个元素即有限空间的情形，特别是对于有限博弈来说，重复剔除的占优战略均衡却是纳什均衡的一种特例。所以我们有：

定理 3.5(蒲勇健，2001)　有限博弈的重复剔除的占优战略均衡是纳什均衡

证明：这是因为，如果某一重复剔除的占优战略均衡 $s^* = (s_1^*, \cdots, s_n^*)$ 不是纳什均衡，则至少存在某一局中人的重复剔除的占优战略均衡战略不是最优战略，不妨设这个局中人就是局中人 1，则存在某个 $s_{11} \in S_1$，使 $u_1(s_{11}, s_2^*, \cdots, s_n^*) > u_1(s_1^*, s_2^*, \cdots, s_n^*)$ 且 s_{11} 被剔除，因为在任一轮剔除中 s_1^* 并不占优于 s_{11}，故必存在某 $s_{12} \in S_1$，s_{12}（严格）在某一轮剔除中占优于 s_{11} 且 $s_{12} \neq s_1^*$，因为 s_{-1}^* 未被剔除，于是有 $u_1(s_{12}, s_2^*, \cdots, s_n^*) > u_1(s_{11}, s_2^*, \cdots, s_n^*) > u_1(s_1^*, s_2^*, \cdots, s_n^*)$，根据同样的推理，我们将得到一个由无穷个战略元素构成的集合 $\pi_1 = \{s_{1n} | s_{1n} \in S_1, n = 1, 2, \cdots\}$，满足

$u_1(s_{1n}, s_2^*, \cdots, s_n^*) > u_1(s_1^*, s_2^*, \cdots, s_n^*)$，$\forall s_{1n} \in \pi_1$

这与假定局中人的战略空间是有限空间的假设相矛盾。

证毕！

冯•诺依曼的最大最小均衡与纳什均衡之间的关系又怎么样呢？我们通过一个例子来说明最大最小均衡不一定就是纳什均衡，即使在二人零和博弈情况下结论也如此。一般地，纳什均衡也不一定是最大最小均衡。

例 3.24　在表 3.24 中，我们给出一个战略式表述的二人零和有限博弈。显然，该博弈的最大最小均衡为(H, L)，但它不是纳什均衡，因为该博弈不存在纯战略纳什均衡[20]。

表 3.24　最大最小均衡不是纳什均衡的二人零和博弈

		B		
		L	M	R
A	U	5, –5	2, –2	3, –3
	H	4, –4	7, –7	10, –10
	D	–1, 1	–3, 3	5, –5

在例 3.25 中，我们给出一个战略式表述的二人有限博弈，其中纯战略纳什均衡为(U, R)，但最大最小均衡为(D, R)。这说明即使是存在纯战略纳什均衡的博弈，最大最小均衡也不一定就是纳什均衡。

例 3.25

最大最小均衡不一定就是纳什均衡，即使博弈存在纯战略纳什均衡，见表 3.25。

表 3.25　最大最小均衡不是纯战略纳什均衡的博弈

		2	
		L	R
1	U	1, 2	3, 4
	D	3, 4	2, 6

纳什均衡也不一定就是最大最小均衡。在例 3.25 中，(U, R)为纳什均衡，但最大最小均衡为(D, R)。尽管如此，对于二人零和博弈，纳什均衡却是最大最小均衡，于是我们有：

定理 3.6　二人零和博弈的纳什均衡是最大最小均衡。

证明：令该博弈为$G=\{S_1,S_2;u_1,u_2\}$且$u_1+u_2=0$，则有

$$\max_{s_2\in S_2}\min_{s_1\in S_1}u_2(s_1,s_2)=\max_{s_2\in S_2}\min_{s_1\in S_1}[-u_1(s_1,s_2)]=\max_{s_2\in S_2}[-\max_{s_1\in S_1}u_1(s_1,s_2)]$$
$$=-\min_{s_2\in S_2}\max_{s_1\in S_1}u_1(s_1,s_2)$$

设$s^*=(s_1^*,s_2^*)$是G的一个纳什均衡，那么对所有$s_2\in S_2$有

$$u_2(s_1^*,s_2^*)\geq u_2(s_1^*,s_2)$$

或者，既然$u_2=-u_1$，则对所有$s_2\in S_2$有$u_1(s_1^*,s_2^*)\leq u_1(s_1^*,s_2)$，因此，$u_1(s_1^*,s_2^*)=\min_{s_2\in S_2}u_1(s_1^*,s_2)\leq\max_{s_1\in S_1}\min_{s_2\in S_2}u_1(s_1,s_2)$。同样地，对所有$s_1\in S_1$，$u_1(s_1^*,s_2^*)\geq u_2(s_1,s_2^*)$，从而对所有$s_1\in S_1$，$u_1(s_1^*,s_2^*)\geq\min_{s_2\in S_2}u_1(s_1,s_2)$，于是$u_1(s_1^*,s_2^*)\geq\max_{s_1\in S_1}\min_{s_2\in S_2}u_1(s_1,s_2)$。这样，$u_1(s_1^*,s_2^*)=\max_{s_1\in S_1}\min_{s_2\in S_2}u_1(s_1,s_2)$并且$s_1^*$是局中人 1 的最大最小化战略。

对于局中人 2 可类似证明，s_2^*是局中人 2 的最大最小战略，并且

$u_2(s_1^*,s_2^*)=\max_{s_2\in S_2}\min_{s_1\in S_1}u_2(s_1,s_2)$。

证毕！

进一步，我们还可证明一个具有纳什均衡的二人零和博弈的所有纳什均衡产生同一支付这一强结论。在非零和博弈中，纳什均衡的这个特性是很少能得到满足的（例如，在例 3.21 给出的斗鸡博弈中，在两个纳什均衡处，局中人的支付分别为 2 和 0）。

推论 3.6.1　如果二人零和博弈存在多个纳什均衡，则所有纳什均衡将给局中人带来相同的支付。

证明：在证明定理 3.6 的过程中已得到：若$s^*=(s_1^*,s_2^*)$为二人零和博弈$G=\{S_1,S_2;u_1,u_2\}$的纳什均衡，则有

$u_1(s_1^*,s_2^*)=\max_{s_1\in S_1}\min_{s_2\in S_2}u_1(s_1,s_2)$和$u_2(s_1^*,s_2^*)=\max_{s_2\in S_2}\min_{s_1\in S_1}u_2(s_1,s_2)$

这说明所有纳什均衡产生相同的支付。

证毕！

一般地，在纳什均衡s^*处不一定有

$$u_1(s^*)=\max_{s_1\in S_1}\min_{s_2\in S_2}u_1(s_1,s_2)=\min_{s_2\in S_2}\max_{s_1\in S_1}u_1(s_1,s_2)$$

但在二人零和博弈下该式却成立。

推论 3.6.2　若s^*是二人零和博弈G的纳什均衡，则有

$u_1(s^*)=\max_{s_1\in S_1}\min_{s_2\in S_2}u_1(s_1,s_2)=\min_{s_2\in S_2}\max_{s_1\in S_1}u_1(s_1,s_2)$。

证明：定理 3.6 的证明过程中已得到

$u_1(s^*)=\max_{s_1\in S_1}\min_{s_2\in S_2}u_1(s_1,s_2)$，　　$u_2(s^*)=\max_{s_2\in S_2}\min_{s_1\in S_1}u_2(s_1,s_2)$

于是

$$u_1(s^*)=-u_2(s^*)=-\max_{s_2\in S_2}\min_{s_1\in S_1}u_2(s_1,s_2)$$
$$=\min_{s_2\in S_2}\max_{s_1\in S_1}[-u_2(s_1,s_2)]=\min_{s_1\in S_2}\max_{s_1\in S_1}u_1(s_1,s_2)$$

所以有

$u_1(s^*)=\max_{s_1\in S_1}\min_{s_2\in S_2}u_1(s_1,s_2)\ =\min_{s_2\in S_2}\max_{s_1\in S_1}u_1(s_1,s_2)$。

证毕！

推论 3.6.2 在某种意义上的逆命题也是成立的，它就是下面的推论 3.6.3。

推论 3.6.3　二人零和博弈 $G=\{S_1,S_2;u_1,u_2\}$ 若满足 $\max_{s_1\in S_1}\min_{s_2\in S_2}u_1(s_1,s_2)$ $=\min_{s_2\in S_2}\max_{s_1\in S_1}u_1(s_1,s_2)$（特别地，如果 G 有一个纳什均衡（参见推论 3.6.2）)，且 s_1^* 是局中人 1 的最大最小战略，s_2^* 是局中人 2 的最大最小战略，则 $s^*=(s_1^*,s_2^*)$ 是 G 的一个纳什均衡。

证明：令 $v^*=\max_{s_1\in S_1}\min_{s_2\in S_2}u_1(s_1,s_2)=\min_{s_2\in S_2}\max_{s_1\in S_1}u_1(s_1,s_2)$，则有

$-v^*=\max_{s_2\in S_2}\min_{s_1\in S_1}[-u_1(s_1,s_2)]=\max_{s_2\in S_2}\min_{s_1\in S_1}u_2(s_1,s_2)$

因 s_1^* 是局中人 1 的最大最小战略，我们有：对所有 $s_2\in S_2$，$u_1(s_1^*,s_2)\ge v^*$；因为 s_2^* 是局中人 2 的最大最小战略，我们有：对所有 $s_1\in S_1$，$u_2(s_1,s_2^*)\ge -v^*$。在第一个不等式中令 $s_2=s_2^*$，我们得到 $u_1(s_1^*,s_2^*)\ge v^*$，于是 $u_1(s_1^*,s_2^*)\ge v^*\ge -u_2(s_1,s_2^*)=u_1(s_1,s_2^*)$；再在第二个不等式中令 $s_1=s^*$，则 $u_2(s_1^*,s_2^*)\ge -v^*\ge -u_1(s_1^*,s_2)=u_2(s_1^*,s_2)$，所以 s^* 是 G 的一个纳什均衡。

证毕！

由推论 3.6.3 可注意到一个纳什均衡可通过求解问题 $\max_{s_1\in S_1}\min_{s_2\in S_2}u_1(s_1,s_2)$ 找到，当计算一个博弈的纳什均衡，特别是当计算混合战略均衡时，这个事实是非常有用的（但通过运用该方法不一定可找出所有混合战略均衡，我们在定理 3.3 中给出的方法尽管带来较大量的计算，但却可保证可找出所有的混合战略均衡，并且该方法还不仅限于零和博弈）。

从定理 3.6 和推论 3.6.3 也可注意到一个二人零和博弈的纳什均衡是可互换的：即如果 (s_1^*,s_2^*) 和 $(\bar{s}_1,\bar{s}_2)$ 是纳什均衡解，则 $(s_1^*,\bar{s}_2)$ 和 $(\bar{s}_1,s_2^*)$ 也是纳什均衡解。

显然，占优战略均衡必定是纳什均衡，因为局中人的占优战略是对所有其他局中人的任何战略组合情况下的最优战略选择，自然它也一定是对于所有其他局中人的某个特定战略（其他局中人的最优战略）情况下的最优战略选择。但反过来不成立，纳什均衡不一定是占优战略均衡，如例 3.13 中的纯战略纳什均衡为(U, R)，但 U 不是局中人 1 的

占优战略（尽管 R 是局中人的占优战略，这也仅仅是一种巧合）。

至此，我们相继引入了几种不同的均衡概念，即最大最小均衡（MME），占优战略均衡（DSE），重复剔除的占优战略均衡（IEDE），纯战略纳什均衡（PNE）和混合战略纳什均衡（MNE）五种均衡概念。我们现在将它们之间的关系用图形集合来表示，见图 3.4。

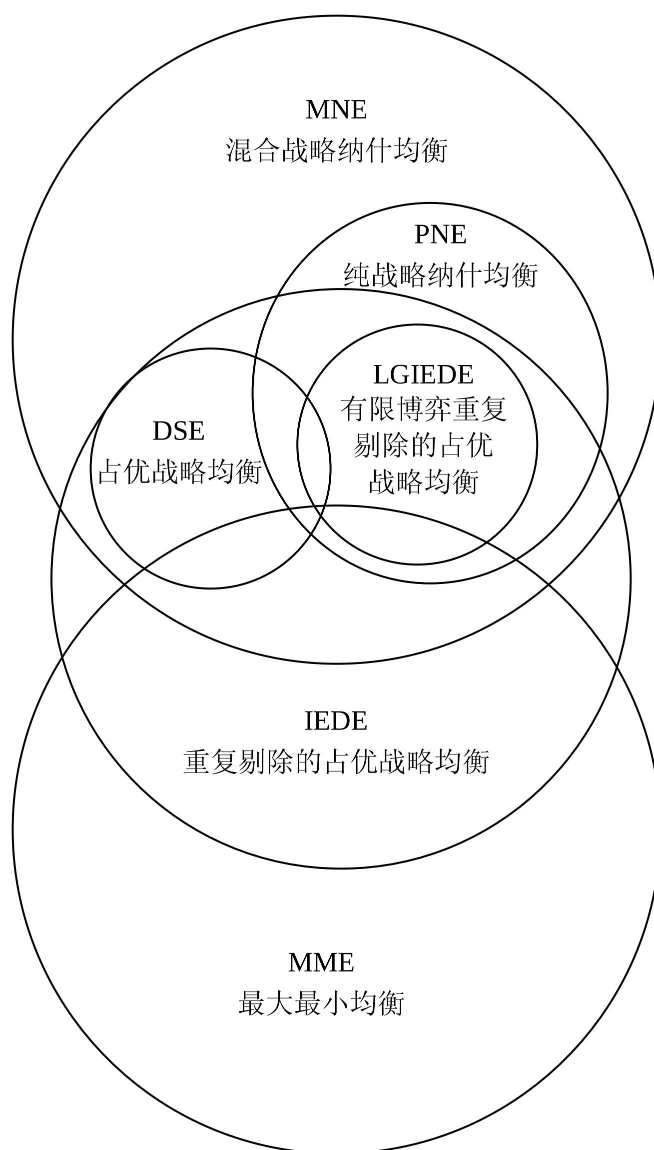


图 3.4　几种均衡概念之间的关系

3.4 完全信息静态博弈：应用与例子

本节将介绍运用完全信息静态博弈模型方法进行经济分析的例子，其中一些例子实际上是经济学中的基本原理，我们将重新从博弈论角度审视这些基本经济学原理的内

涵，这样有助于加深对一些经济学现象及原理的理解。另一些是学者们运用博弈论在经济学中获得的发现，这些发现证实了博弈论作为经济分析工具的神奇功能。这类例子不仅在经济学中具有重要的意义，而且对于研究者来说还具有方法论上的启发性。在进行这种介绍之前，我们先来看看如何运用纳什均衡概念重新处理例 3.10 所给出的古诺寡头竞争模型。在例 3.10 中，我们运用重复剔除劣战略的方法得到了重复剔除的占优战略均衡，下面，我们将证明，这个均衡还是纳什均衡，尽管这种博弈并不是有限博弈。

例 3.26　古诺寡头竞争模型：纳什均衡

在例 3.10 中，我们已给出了局中人 i 的支付函数为

$$\pi_i(q_i,q_2)=q_i p(q_1+q_2)-c_i(q_i)$$

其中 q_i 为 i 的产量， $p(q_1+q_2)$ 为逆需求函数， p 为产品价格， $i=1,2$

给定 q_j ， i 的最优产量 q_i 满足一阶条件：

$$\frac{\partial\pi_i}{\partial q_i}=p(q_1+q_2)+q_i p'(q_1+q_2)-c'_i(q_i)=0,\quad i=1,2$$

反应函数为 $q_i^*=R_i(q_j)$ ， $i\neq j$ ， $i,j=1,2$ 。

假定局中人 i 的成本函数为 $c_i(q_i)=cq_i, i=1,2$，其中 c 为常数平均成本，且假设逆需求函数为线性的：

$$p=a-(q_i+q_2)$$

于是一阶条件变为： $a-(q_1+q_2)-q_i-c=0,\quad i=1,2$

反应函数为

$$q_i^*=\frac{1}{2}(a-q_j-c),\quad i\neq j,\ i,j=1,2$$

由对称性，必有

$$q_i^*=\frac{1}{2}(a-q_i^*-c),\quad i=1,2$$

得

$$q_i^*=\frac{1}{3}(a-c),\ i=1,2$$

每个企业的利润为

$$\pi_i=\left[a-2\times\frac{1}{3}(a-c)\right]\cdot\frac{1}{3}(a-c)-c\cdot\frac{1}{3}(a-c)=\frac{1}{9}(a-c)^2$$

为了与垄断情况进行比较，让我们再来计算一下当两个企业联手进行完全垄断经营时的最优产量和利润。在垄断情形，两个企业最大化总的垄断利润，表达为：

$$\max_Q \pi_i=Q(a-Q-c)$$

其中 Q 为两个企业的产量之和。

解得

$$Q^* = \frac{1}{2}(a-c) < q_1^* + q_2^* = \frac{2}{3}(a-c)$$

而垄断利润为$\frac{1}{4}(a-c)^2 > \pi_1 + \pi_2 = \frac{2}{9}(a-c)^2$。

寡头竞争的总产量大于垄断产量的原因在于每个企业在选择自己的最优产量时，只考虑对本企业利润的影响，而忽视对另一个企业的外部负效应。这本质上仍是一个囚徒困境问题。从经济学上看，寡头竞争比起完全垄断来说仍然含有一定的竞争性，因而产量比完全垄断产量要大一些，而利润（价格）又比完全垄断情形低一些（完全垄断价格为$\frac{1}{2}(a-c)$，而寡头垄断价格为$\frac{1}{3}(a-c)$）。在垄断情况，两个企业的总利润大于寡头竞争时的总利润，这似乎意味着两个企业会联手串谋进行统一定价对市场进行垄断经营，然后再平分利润。但是，由于两个企业之间处于一种囚徒困境，串谋合作通常是不稳定的。给定另一个企业遵守协议，另一个企业就有动机违反协议。当然，它们之间似乎可以通过订立协议来约束破坏合作的行为，但这种公开协议通常是违法的。但即使可以订立公开的协议，企业仍然可能私下违反协议，如国际石油卡特尔 OPEC 经常就成员国违反协议的情况进行协调，但效果不能持久，成员国违反协议的行为屡屡出现，据说 20 世纪 90 年代初伊拉克入侵科威特时就打着惩罚科威特曾多次违反协议的幌子。因此，在描述寡头行为时，经济学家通常认为古诺模型是一个比垄断模型更加可取的方法。

例 3.27　公共地的悲剧

经济学家很早就观察到一个现象，当资源产权没有得到明晰界定时，资源就会被过度利用。在 18 世纪以前，苏格兰有着广褒的草地和牧场，但到了 19 世纪，这些天然大牧场再也见不到了。由于草场的所有权没有界定，牧民们的过度放牧使草场失去了再生能力，牧草走向自然耗竭。类似的现象还发生在其它公共资源耗竭的情形，如公海中天然鱼类的灭绝；中国农村小煤窑的过度泛滥导致自然资源过度开发、环境污染；中国长江上游天然森林的过度砍伐（它是导致自 20 世纪 90 年代以来长江中下游连年洪水泛滥的原因，因为长江上游森林植被遭到破坏导致水土流失，在中下游积成大洪水）等等。对于这类现象的关注早在 18 世纪的学者们那里就有所著述，至迟从休漠（1739）开始，政治哲学和经济学家已经认识到如果公民只关注个人福利，公共物品就会出现短缺，并且公共资源也会过度使用。今天，只要随便看一下地球的环境，就能体会到这一观念的力量。在这里给出的例子里，我们以牧场的过度利用作为模型来说明其中的道理。

现假定有n个牧民共有一片草地，每个牧民可在草地上自由地放羊。$g_i \in [0,\infty)$是第i个农民所养的羊数量，$i = 1,\cdots,n$。$G = \sum_{i=1}^{n} g_i$是总的羊数量，v表示每只羊的平均价值。因为草地面积固定，故羊愈多，每只羊获取的营养就愈小，其价值就愈低（体重愈低），故$V = V(G)$。

每只羊至少要有在一定数量下限以上的草料供养才能存活，故存在一个最大可存活的羊数量$G_{\max}$：

当$G < G_{\max}$，$V(G) > 0$；

当$G \geq G_{\max}$，$V(G) = 0$。

因羊数量愈多，V 就愈小。故$\frac{\partial V}{\partial G} < 0$；当羊数量很小时，羊数量增加导致羊之间在草料上的竞争不太激烈，故$\frac{\partial V}{\partial G}$的绝对值较小；但当$G$较大时，羊之间争夺草料较激烈，$\frac{\partial V}{\partial G}$的绝对值就较大，故$\frac{\partial^2 V}{\partial G^2} < 0$（即草料对羊的价值的边际产出递减），图 3.5 给出了$V(G)$的图象。

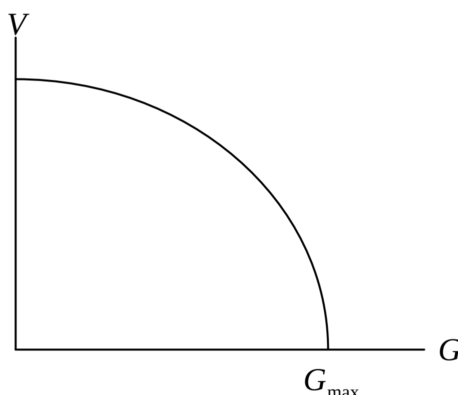


图 3.5　$V(G)$的图象

牧民i的战略空间为$S_i = \{g_i | g_i \in [0, \infty)\}$，设购买和照看一只羊的成本为$c$，则每个牧民的利润函数为

$$\pi_i(g_1, \cdots, g_n) = g_i V(G) - g_i c,\ i = 1, \cdots, n$$

一阶条件为：

$$\frac{\partial V}{\partial g_i} = V(G) + g_i V'(G) - c = 0,\ i = 1, \cdots, n \quad （3.40）$$

n个方程可解出n个未知数$g_i^*(i = 1, \cdots, n)$，它们就是纳什均衡。不难证明该均衡是低效率的，即每个牧民养的羊太多（这也是一种囚徒困境）。

将一阶条件（3.40）中的n个方程相加：

$$nV(G^*) + G^* V'(G^*) - nc = 0 \quad （3.41）$$

再在式（3.41）两端同除以n：

$$V(G^*) + \frac{G^*}{n} V'(G^*) = c\text{，其中 } G^* = \sum_{i=1}^{n} g_i^* \quad （3.42）$$

如果n个牧民联合行动，最大化总利润$\pi = GV(G) - Gc$，则需解下述帕累托最优问

题：

$$\max_G [GV(G)-Gc]$$

一阶条件为：

$$V(G^{**})+G^{**}V'(G^{**})=c \quad (3.43)$$

将式（3.42）减去式（3.43），再移项得：

$$V(G^*)-V(G^{**})=G^{**}V'(G^{**})-\frac{G^*}{n}V'(G^*) \quad (3.44)$$

则必有$G^*>G^{**}$，若不然，设$G^*\le G^{**}$，因$\frac{\partial V}{\partial G}<0$，故

$V(G^*)-V(G^{**})\ge 0$，

由式（3.44）有

$$G^{**}V'(G^{**})\ge \frac{G^*}{n}V'(G^*) \quad (3.45)$$

又因$\frac{\partial^2 V}{\partial G^2}<0$，故

$$V'(G^*)\ge V'(G^{**})$$

但$V'(G^*)<0$，故

$$0<-V'(G^*)\le -V'(G^{**}) \quad (3.46)$$

由不等式（3.45）有：

$$G^{**}\left[-V'\left(G^{**}\right)\right]\le \frac{G^*}{n}\left[-V'\left(G^*\right)\right]$$

由不等式（3.46），必有：

$$G^{**}\le \frac{G^*}{n}$$

即$G^*\ge nG^{**}>G^*(n>1)$，这与原假设$G^*\le G^{**}$矛盾。所以必有$G^*>G^{**}$，即公有草地被过度利用了。

例 3.28　公共品的私人供给不足。

与公共资源的过度利用形成对照的是公共物品的私人供给不足。在经济学中我们已熟知公共品的私人供给存在“外溢”效应，即所谓外部性。在这种情形，如果公共品由私人供给，则存在与帕累托最优相比较的供给不足问题。孟德维尔早在《蜜蜂的寓言》中就以果园主与养蜂人的故事向我们讲述了这种供给不足的道理。根据孟德维尔的理论，果园中果树盛开的芬芳花朵为密蜂提供了采蜜的环境而增加了养蜂人的收益，同时果树又由于蜜蜂的授粉而增加了果园主的收益，但由于果园主并未向养蜂人支付价格而使养蜂人养的蜜蜂太少，反过来又由于养蜂人并未向果园主支付价格又使果园主种的果树太少。因此，孟德维尔建议果园主要向养蜂人支付价格，养蜂人也应向果园主支付价格，这样才可能使果树的种植数和蜜蜂的饲养量达到帕累托最优。

下面，我们用居民自愿修筑一个防洪大坝为例说明公共品私人供给不足的道理。

假定由n个居民建筑一座防洪大坝，每个居民自愿提供沙袋。第i个居民的贡献为g_i，总供给为$G=\sum_{i=1}^{n} g_i$。第i个居民的效用函数为$u_i(x_i,G)$，x_i是居民i的其他私人物品消费量，$i=1,\cdots,n$。假定$\frac{\partial u_i}{\partial x_i}>0$，$\frac{\partial u_i}{\partial G}>0$（经济学中关于边际效用为正的通常假定），令$P_x$为$x_i$的价格，$P_G$为沙袋价格，$M_i$为个人总预算收入。居民总是在预算约束下最大化个人效用：

$$\begin{cases}\max\limits_{x_i,g_i} u_i(x_i,G) \\ s.t.\quad P_x X_i + P_G g_i = M_i\end{cases} \tag{3.47}$$

构造问题（3.47）的拉格朗日函数。

$$L_i(X_i,G,\lambda_i)=u_i(x_i,G)+\lambda_i(M_i-P_x x_i-P_G g_i)$$

其中λ_i是拉格朗日乘子。

一阶条件：

$$\left.\begin{array}{l}\frac{\partial u_i}{\partial G}-\lambda_i P_G=0 \\ \frac{\partial u_i}{\partial x_i}-\lambda_i P_i=0\end{array}\right. ,\quad i=1,\cdots,n \tag{3.48}$$

将（3.48）中的两式相除得：

$$\frac{\frac{\partial u_i}{\partial G}}{\frac{\partial u_i}{\partial x_i}}=\frac{P_G}{P_X} \tag{3.49}$$

式（3.49）就是微观经济学中的“等边际法则”。

式（3.48）中$2n$个方程决定了纳什均衡$g^*=(g_1^*,\cdots,g_n^*)$和拉格朗日乘子$\lambda=(\lambda_1^*,\cdots,\lambda_n^*)$，总供给均衡量为$G^*=\sum_{i=1}^{n} g_i^*$。对整个社会来说，帕累托最优的$G^*$由社会福利函数的最大化条件决定，不妨设社会福利函数为：

$$w=\sum_{i=1}^{n}\gamma_i u_i,\quad \gamma_i\ge 0,\quad \sum_{i=1}^{n}\gamma_i=1 \tag{3.50}$$

社会总预算约束为

$$P_x\sum_{i=1}^{n}x_i+P_G G=\sum_{i=1}^{n}M_i \tag{3.51}$$

拉格朗日函数为

$$L(x_1,\cdots,x_n,G,u)=\sum_{i=1}^{n}\gamma_i u_i+u\left[\sum M_i-P_x\sum_{i=1}^{n}x_i-P_G G\right]$$

一阶条件：

$$\left.\begin{aligned}&\sum_{i=1}^{n}\gamma_i\frac{\partial u_i}{\partial G}-uP_G=0\\&\gamma_i\frac{\partial u_i}{\partial x_i}-uP_x=0\end{aligned}\right.\quad,\quad i=1,\cdots,n \tag{3.52}$$

其中u是拉格朗日乘子。

由（3.52）中后n个方程有：

$$\gamma_i=\frac{uP_x}{\dfrac{\partial u_i}{\partial x_i}}$$

分别代入式（3.52）中的前n个方程：

$$\sum_{i=1}^{n}uP_x\frac{\dfrac{\partial u_i}{\partial G}}{\dfrac{\partial u_i}{\partial x_i}}=uP_G \tag{3.53}$$

得到：

$$\sum_{i=1}^{n}\frac{\dfrac{\partial u_i}{\partial G}}{\dfrac{\partial u_i}{\partial x_i}}=\frac{P_G}{P_x} \tag{3.54}$$

式（3.54）被称为“存在公共品情况下的帕累托最优的萨缪尔逊条件”，它是等边际法则（3.49）在存在公共品情况下的推广。

将条件（3.54）重新写为：

$$\frac{\dfrac{\partial u_j}{\partial G}}{\dfrac{\partial u_j}{\partial x_j}}=\frac{P_G}{P_x}-\sum_{i\neq j}\frac{\dfrac{\partial u_i}{\partial G}}{\dfrac{\partial u_i}{\partial x_i}} \tag{3.55}$$

在经济学中，$\dfrac{\dfrac{\partial u_j}{\partial G}}{\dfrac{\partial u_j}{\partial x_j}}$是$G$对$x_j$的边际替代率，一般假设它是$G$的递减函数。对比式（3.55）与等边际法则（3.49），可知社会最优的公共品供给G^{**}大于纳什均衡公共品供给G^*，即$G^{**}>G^*$，说明私人自愿供给的公共品存在不足。[21]

例 3.29　Bertrand 寡头模型

在例 3.10 和例 3.26 中，我们曾假定寡头的纯战略为产量，这里，我们再来看看寡头的纯战略为价格的情形。寡头间的竞争仍假定为“古诺式”的，即每个企业在给定其它企业价格水平下选择自己的最优价格水平。[22]

假定：两个寡头企业的产品有一定差异并且进行价格竞争。产品有一定差异指两厂

商的产品是在品牌、质量、包装等方面有所不同的同类产品。因此 Bertrand 模型中厂商的产品之间有很强的替代性，但又不是完全可替代，即价格不同时，价格较高的不会完全销不出去。这种情况，我们可用当厂商 1 和厂商 2 的价格分别为 P_1 和 P_2 时，他们各自面临的需求函数分别为

$$q_1 = q_1(p_1, p_2) = a_1 - b_1 p_1 + d_1 p_2, \qquad q_2 = q_2(p_1, p_2) = a_2 - b_2 p_2 + d_2 p_1$$

来表示，其中 $d_1, d_2 > 0$ 表示两厂商产品有一定替代性的替代系数。我们同样假设两厂商无固定成本，边际生产成本分别为 c_1 和 c_2。

在该博弈中，两厂商各自的战略空间分别为 $S_1 = [0, P_{1\max}]$ 和 $S_2 = [0, P_{2\max}]$，其中 $P_{1\max}$ 和 $P_{2\max}$ 分别是厂商 1 和厂商 2 还能卖出产品的最高价格[23]。

两厂商的利润函数为

$$\pi_1(P_1, P_2) = P_1 q_1 - c_1 q_1 = (P_1 - c_1) q_1 = (P_1 - c_1)(a_1 - b_1 P_1 + d_1 P_2)$$

$$\pi_2(P_1, P_2) = P_2 q_2 - c_2 q_2 = (P_2 - c_2) q_2 = (P_2 - c_2)(a_2 - b_2 P_2 + d_2 P_1)$$

由一阶条件得反应函数：

$$P_1 = \frac{1}{2b_1}\left(a_1 + b_1 c_1 + d_1 P_2\right)$$

$$P_2 = \frac{1}{2b_2}\left(a_2 + b_2 c_2 + d_2 P_1\right)$$

设纳什均衡为 $\left(P_1^*, P_2^*\right)$，必有

$$\begin{cases} P_1^* = \dfrac{1}{2b_1}\left(a_1 + b_1 c_1 + d_1 p_2^*\right) \\ P_2^* = \dfrac{1}{2b_2}\left(a_2 + b_2 c_2 + d_2 p_1^*\right) \end{cases}$$

解得：

$$\begin{cases} P_1^* = \dfrac{d_1}{4b_1 b_2 - d_1 d_2}\left(a_2 + b_2 c_2\right) + \dfrac{2b_2}{4b_1 b_2 - d_1 d_2}\left(a_1 + b_1 c_1\right) \\ P_2^* = \dfrac{d_2}{4b_1 b_2 - d_1 d_2}\left(a_1 + b_1 c_1\right) + \dfrac{2b_1}{4b_1 b_2 - d_1 d_2}\left(a_2 + b_2 c_2\right) \end{cases}$$

当两厂商的产品完全无差异时，该模型中的需求函数要修改，此时必须考虑消费者对价格的敏感性。如果所有消费者对价格都非常敏感，则两厂商的价格竞争将导致均衡价格等于边际成本。在这种情形，即使不是完全竞争，价格也等于边际成本，当固定成本为零且边际成本为常数时，利润就等于零，这被称为“Bertrand 悖论”[24]。

需要指出，这种价格决策与古诺模型中的产量决策一样，其纳什均衡也是低效率的，因为通过合作形成完全垄断经营各个局中人可以得到更多的利润，但同样他们也处于囚徒困境之中，这种合作是不稳定的。读者可自行证明此点。

例 3.30　Hotelling 价格竞争模型

有人说倘若 Hotelling 教授还在世的话，他会因为提出著名的 Hotelling 模型而获得诺贝尔经济学奖。在现代经济学（特别是产业组织理论）的许多分支中，Hotelling 模型为我们分析诸多“差异化”问题提供了一个统一并且有效的工具。譬如，在产业组织理论中，我们关心生产同样产品的不同企业之间如何通过在产品或服务上增添一些带有特征性的属性来增强自身的相对垄断性，从而在激烈的竞争中保有一个相对垄断性的市场，此时，我们用 Hotelling 模型就可以作出很好的分析。Hotelling 模型分为价格竞争模型和位置竞争模型。这里，为了简明地介绍该模型的要诣，我们先来看看 Hotelling 价格竞争模型的内容。在后面的例 3.31 中，我们再对位置竞争模型予以介绍。

在完全竞争行业中，企业利润在行业长期均衡时应为零。Bertrand（1883）进一步证明，即使只有两个企业，在均衡时价格也会等于边际成本，如果固定成本为零且边际成本为常数，则企业利润仍为零。因为倘若不同企业的产品对于消费者来说是完全“同质的”，则企业之间的竞争将使价格跌落在边际成本上才达到均衡。这样，即使只有两个生产“同质”产品的企业，价格也会等于边际成本、这与完全竞争市场是一样的。这被称为“Bertrand 悖论”。然而，我们在现实中观察到的现象并非如此，如寡头利润往往很高，一些公司在长期也获得可观的利润。要破译 Bertrand 悖论，就需要放弃关于“同质”产品或服务的假设，引入不同企业在产品或服务上的差异性。当不同企业的产品或服务存在差异性时，不同消费者对不同产品或服务有着不同的偏好，价格不再是他们感兴趣的唯一变量，此时，均衡价格不再等于边际成本，利润就会大于零。在这个例子中，与例 3.26 不同，我们用空间差异性而不是需求函数来刻画产品或服务对需求的影响，但实际上这种方法可用于一般广泛得多的产品或服务差异性处理，如“位置”或空间差异性可代表消费者喜欢软饮料“甜味”的程度。

假定：产品在物质性能上完全相同，但在空间位置上有差异。因不同位置上的消费者要支付不同的运输成本，他们关心的是价格与运输成本之和，而不单是价格。

假设有一长度为 1 的线性城市，消费者在[0, 1]上均匀分布，分布密度为 1。有两个商店分别位于城市两端，商店 1 在 $x=0$，商店 2 在 $x=1$，出售相同产品。每个商店提供单位产品的成本为 c，消费者购买商品的旅行成本与其离商店的距离成正比，比例系数为 t，故在 x 处的消费者若在商店 1 购买，其旅行成本为 tx；若在商店 2 购买，旅行成本为 $t(1-x)$，见图 3.6 所示。

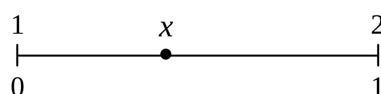


图 3.6　Hotelling 城市

假定消费者具有单位需求，即或者消费 1 个单位或者消费 0 个单位。下面求解两个

商店之间进行价格竞争的纳什均衡。假定两个商店同时选择自己的价格。为简化分析，假定每个消费者都购买一个单位产品（即消费者剩余相对于价格加上旅行费用足够大）。令 P_i 为商店 i 的价格，$D_i(P_1,P_2)$ 为其需求函数，$i=1,2$。

设 x 处的消费者对两个商店之间是无差异的，则在 x 左边的消费者都会在商店 1 购买产品，而 x 右边的消费者会都在商店 2 购买，故需求分别为 $D_1=x$，$D_2=1-x$，x 满足：

$$P_1+tx=P_2+t(1-x)$$

解得：

$$x=\frac{P_2-P_1+t}{2t}$$

$$D_1(P_1,P_2)=x=\frac{P_2-P_1+t}{2t}$$

$$D_2(P_1,P_2)=1-x=\frac{P_1-P_2+t}{2t}$$

利润函数为：

$$\pi_1(P_1,P_2)=(P_1-c)D_1(P_1,P_2)=\frac{1}{2t}(P_1-c)(P_2-P_1+t)$$

$$\pi_2(P_1,P_2)=(P_2-c)D_2(P_1,P_2)=\frac{1}{2t}(P_2-c)(P_1-P_2+t)$$

商店 i 选 P_i 最大化 π_i，一阶条件为：

$$\frac{\partial\pi_1}{\partial P_1}=P_2+c+t-2P_1=0$$

$$\frac{\partial\pi_2}{\partial P_2}=P_1+c+t-2P_2=0$$

得均衡解：

$$P_1^*=P_2^*=c+t$$

利润为：

$$\pi_1=\pi_2=\frac{t}{2}$$

如果将位置差异 t 解释为产品差异，则差异愈大，均衡价格及利润就愈高。如果 $t=0$，则回到 Bertrand 悖论，价格等于边际成本，利润为零。这个模型对寡头企业存在大于零的长期利润的解释就是：寡头长期利润来自于其产品或服务的差异性。

例 3.31　Hotelling 位置竞争模型

在例 3.30 中，我们给定厂商的位置而确定厂商的价格。在本例中，我们改变这个假定，假设厂商的价格给定条件下如何通过调整厂商的位置来进行竞争[25]。

为了说明这个模型的均衡，我们假定 Hotelling 的线性城市是以离散方式测度其位置的，或者说假定商店总是占有一个单位的地理范围。在图 3.7 中，Hotelling 城市可有

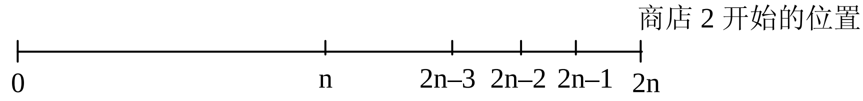


图 3.7　Hotelling 城市

$2n+1$个位置可供一个商店选择（n为充分大的自然数）。开始，假定商店 2 的位置固定在$2n$处，我们来看看商店 1 应该如何选择其最优位置。与例 3.30 一样，我们假定两个商店出售的产品是“同质”的，除了消费者与商店之间的地理距离不同外，消费者对两个商店出售的产品一视同仁，但消费者每旅行一个单位路程的成本仍为t。与例 3.30 不同的假设是，我们这里假定两个商店出售产品的价格是相同的，譬如说因政府对价格进行了规定。显然，给定商店 2 的初始位置，商店 1 的最优选址是紧靠商店 2 的左端即在$2n-1$处[26]，这样可以使除$2n$处的那个消费者外所有其他消费者离商店 1 的距离比离商店 2 的距离小。此时，商店 1 面临的需求为$2ne$，而商店 2 面临的需求为$1\cdot e$，假设n充分大时，则有$2ne>e$[27]，其中$e>0$为比例系数。但是，给定商店 1 选址于$(2n-1)$处，商店 2 原有的选址$2n$不是其最优选址，它也会调整其位置。显然，此时商店 2 的最优选址为$2n-2$（当n充分大时，此时商店 2 面临的需求$(2n-1)e>e$）。同样，给定商店 2 选址于$(2n-2)$，商店 1 前一次选定的位置$(2n-1)$又不再是其最优选址，此时它的最优选址应为$(2n-3)$（因当n充分大时，商店 1 位于$(2n-3)$面临的需求为$(2n-2)e>$商店 1 位于$(2n-1)$处面临的需求$2e$）；显然，这种调整互动过程会一直进行下去，直到某一个商店选址于n，而另一个商店选址于$n+1$或$n-1$（此时两种选择对这个商店来说是无差异的）。这样，我们可得到两个纳什均衡。如果两个商店按前面给出的顺序进行位置调整，则当n为偶数时，商店 2 的均衡位置为n，商店 1 的均衡位置为$n+1$或$n-1$；当n为奇数时，商店 1 的均衡位置为n，商店 2 的均衡位置为$n+1$或$n-1$。

但是，我们没有理由假定两个商店的选址决策过程一定是按上述顺序进行的。从静态博弈看，两个商店都同时在进行博弈决策。下面，我们运用反应函数方法来找出纳什均衡点。设商店i的位置为s_i，战略空间为$S_i=\{0,1,\cdots,2n\}$，$s_i\in S_i$，$i=1,2$。

给定s_j，i的最优选择为

$$s_i=\begin{cases} s_j-1 & s_j>n \\ n+1\text{或}n-1 & s_j=n,\ i\neq j,\ i,j=1,2 \\ s_j+1 & s_j<n \end{cases} \tag{3.56}$$

式（3.56）给出了两个博弈局中人的反应函数，见图 3.8。

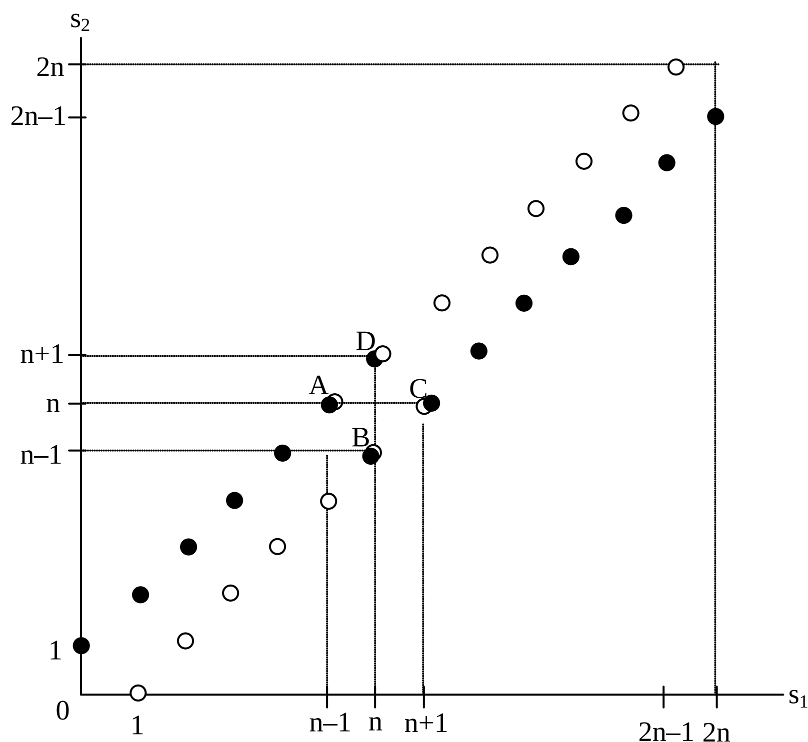


图 3.8　两个商店的反应函数

实心圆点表示商店 2 的反应函数，空心圆点表示商店 1 的反应函数

注意，在图 3.8 中，A、B、C、D 四个位置上的实心和空心圆点实际上应位于同样的空间位置，这里将它们分开只是因为不可能将实心和空心的圆点画在同一空间位置处。空心圆点表示商店 1 的反应函数，实心圆点表示商店 2 的反应函数。两个反应函数的交点 A、B、C、D 就是四个纳什均衡，它们分别为：

A：(n–1, n)，B：(n, n–1)，C：(n+1, n)，D：(n, n+1)

在这个例子中，我们假定 Hotelling 城市中可作为商店选址的位置有奇数个，但这个假设不具有特别性：因为我们总是可以将[0, 1]线段划分为奇数个子区间。另外，即使是偶数个位置，类似的分析也是可以进行的。

该模型的一个特点是：均衡选址都是两个商店毗邻的。这个例子解释了这样一种观察到的现象，即不少市场中都有厂商位置互相邻近的倾向。例如，通常有几家加油站位于繁忙十字路口的拐角处。

例 3.30 和例 3.31 分别考虑了价格可调整而位置固定和位置可调整而价格固定的竞争均衡情形，接下来一个十分自然的引伸考虑就是能否同时假定价格和位置都可调整的情形。不幸的是，当厂商可无成本地改变它们的价格和位置（例如，重新设计它们的产

品）时，可证明纯战略均衡不存在（D' Aspremont, Gabszewicz 和 Thisse, 1979）[28]。但是，这个模型经修改后仍存在纯战略均衡。一种修改法是允许非线性交通成本的存在，另一种方法将在下面的例 3.32 中给出。

例 3.32　塞洛普圆周模型

一些模型修改了 Hotelling 的基本模型从而使均衡存在。这些模型中最有趣最知名的模型之一是塞洛普（Salop, 1979）提出的圆周模型，这个模型对 Hotelling 模型作了两处大的修改。

第一，在这个模型中，厂商位于一个圆周上而不是在一条直线上。作出这一修改的原因是圆周没有终点。即，在没有终点这一点上，圆周与一条无限长的直线是大致相同的。在 Hotelling 的价格和位置同时可调整的模型中纯战略均衡不存在的一个主要原因是终点的存在。

第二，塞洛普模型明确考虑了第二产品或外部商品。例如，差别产品可能是冰淇淋（圆周上的产品）的品牌（风味），而外部产品可能是由另一行业竞争性供给的无差别产品——巧克力蛋糕。

①消费者行为

假设顾客均匀地分布于单位圆周上。为简单起见，每位顾客购买的冰淇淋恰为一勺。顾客的位置 t^* 表示顾客最偏好的冰淇淋类型。譬如说，假设圆周上一个位置是巧克力冰淇淋，另一个位置是香草冰淇淋，而巧克力冰淇淋与香草冰淇淋之间的一点是巧克力条冰淇淋。每种风味的冰淇淋都是一种可能的品牌并通过它在圆周上的位置被加以描述。

假定消费者从食用一勺位于 t 位置的某种冰淇淋所获得的愉悦（效用）用如下函数刻画：

$$u(t,t^*) = u - c\,|\,t - t^*\,| \tag{3.57}$$

其中 u 是消费者从他最喜爱的风味的冰淇淋中获得的效用（这一风味的冰淇淋位于与消费者相同的点——沿圆周的 t^* 处）；$|\,t - t^*\,|$（t 与 t^* 之间距离的绝对值）是品牌 t 与顾客最喜爱的风味 t^* 之间的距离；c 是与最优品牌的偏差降低消费者愉悦的比例系数。

消费者效用函数见图 3.9，图中部分圆周已被拉直成为一条直线。

该图说明在 $t = t^* + u/c$ 和 $t = t^* - u/c$ 处，消费者效用为零。该图还说明消费者从一个位于最优品牌左侧或右侧的品牌中获得的愉悦，低于从最优品牌中获得的愉悦。

每位消费者都试图使消费者剩余最大化，消费者剩余即消费者食用位于 t 处的品牌获得的愉悦与价格之间的差额为 $u(t,t^*) - p$。换种说法，如果你最喜爱的冰淇淋风味是巧克力，但巧克力条冰淇淋的价格只有巧克力冰淇淋的一半，你就可能购买巧克力条冰淇淋，因为你在味道或效用上的损失低于你购买较便宜产品获得的收益。因此，你选择最佳购买：具有最大剩余的产品——价格与质量的最佳组合。

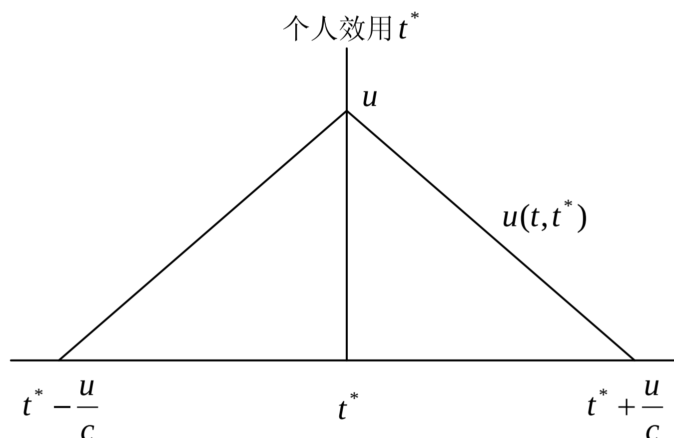


图 3.9　消费者效用

但是，消费者可能决定购买外部商品巧克力蛋糕，而不是某一品牌的冰淇淋，如果从给定数量的货币下外部产品给予消费者更多愉悦的意义上说，这是一种更佳选择。假定蛋糕产生的剩余(吃蛋糕产生的愉快减去蛋糕的价格)为 $\underline{u}$ 。如果一勺最佳购买品牌的冰淇淋 i 的剩余超过 $\underline{u}$ ，消费者将只购买这种冰淇淋：

$$\max_i[u(t_i,t^*)-p_i]\geq \underline{u} \tag{3.58}$$

不等式左边是最佳购买品牌冰淇淋产生的剩余（通过选择品牌 i 使剩余最大化），而不等式右边是蛋糕产生的剩余。就是说，如果最佳购买品牌的冰淇淋产生的剩余至少与蛋糕产生的剩余一样大时，消费者应只购买冰淇淋。

消费者可以从最佳购买冰淇淋中获得的最大剩余是 $u-p^*$，这里 p^* 是位于 t^* 处的冰淇淋品牌价格。只有当品牌剩余大于蛋糕产生的剩余时即 $u-p^*\geq\underline{u}$ 或者 $u-\underline{u}\geq p^*$，消费者才愿意购买那一品牌。其结果是，消费者有一保留价格即消费者愿意为那一冰淇淋品牌所支付的最高价： $v=u-\underline{u}$ 。

换种说法，只有当最佳购买品牌产生的净剩余即最佳购买品牌产生的剩余减去蛋糕产生的剩余为正时，消费者才会购买这一勺冰淇淋：

$$\max_i[v-c\,|\,t_i-t^*|-p_i]\geq 0 \tag{3.59}$$

不等式（3.59）是通过将不等式（3.58）两边减去 $\underline{u}$ ，根据方程（3.57）将 $u(t,t^*)$ 替换，并使用 $v=u-\underline{u}$ 后得到的。

②厂商行为

这一模型的对称均衡取决于厂商位于何处及厂商如何定价。其它保持不变，每家厂商都想在距离它最近的竞争者尽可能远的地方定位。其他商店距离你的商店愈远，对位于你商店附近的顾客而言你的市场力量愈大。由于各商店试图离得尽可能远，它们在互

相等距离处定位。如果圆周上等距离地有 n 个冰淇淋品牌，两品牌间的距离便是 $\frac{1}{n}$（因为圆周为单位圆周）。

塞洛普通过假设各商店已在互相等距离处定位而开始他的分析，然后提出每家商店索取何种价格的问题。假设一典型品牌索价为 p，它最邻近的两个竞争者索价为 $\underline{p}$，如图 3.10 所示。典型品牌的生产者应如何确定 p？答案取决于存在多少种品牌。我们首先考虑存在相对较少厂商的情况，然后考虑具有多得多厂商的市场。

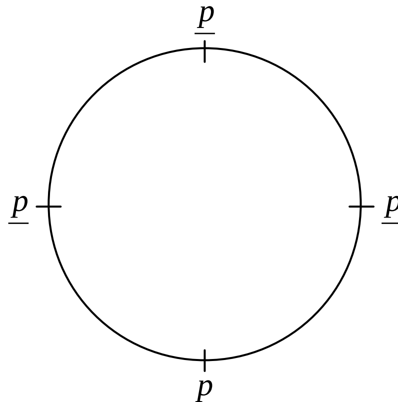


图 3.10　圆周市场

③垄断区域

如果品牌相对较少，它们就不会为了同一顾客群而互相竞争。每一品牌都是一个区域性垄断厂商并向住得足够近的所有消费者出售产品以使他们的净剩余为正。意即每一垄断厂商仅向那些从它的品牌中获取比蛋糕所产生的剩余更大的剩余的消费者出售产品。

考虑一位与位于 t 处价格为 p 的品牌的距离为 $x=|t-t^*|$ 的消费者。只有当消费者的净剩余为正时，即 $v-cx-p\geq 0$（使用条件（3.59）中剩余的表达式），消费者才愿意购买这一品牌。于是通过表达式重组，可得出消费者仍购买这一品牌的离品牌位置的最大距离为：

$$x_m=\frac{v-p}{c} \tag{3.60}$$

距离 x 的决定如图 3.11 所示。图中的纵轴表示这一品牌产生的净剩余；横轴表示消费者离开最偏好品牌（标价为 p 并假设 p 略高于 $\underline{p}$）的距离 x。一个品牌离开消费者最偏好产品的距离愈远，消费者的净剩余愈低。当品牌距离消费者最偏好品牌的位置为 x_m 时，消费者从这一品牌获得的净剩余等于零（净剩余线落至 x 轴），所以消费者不在乎购买或不购买。

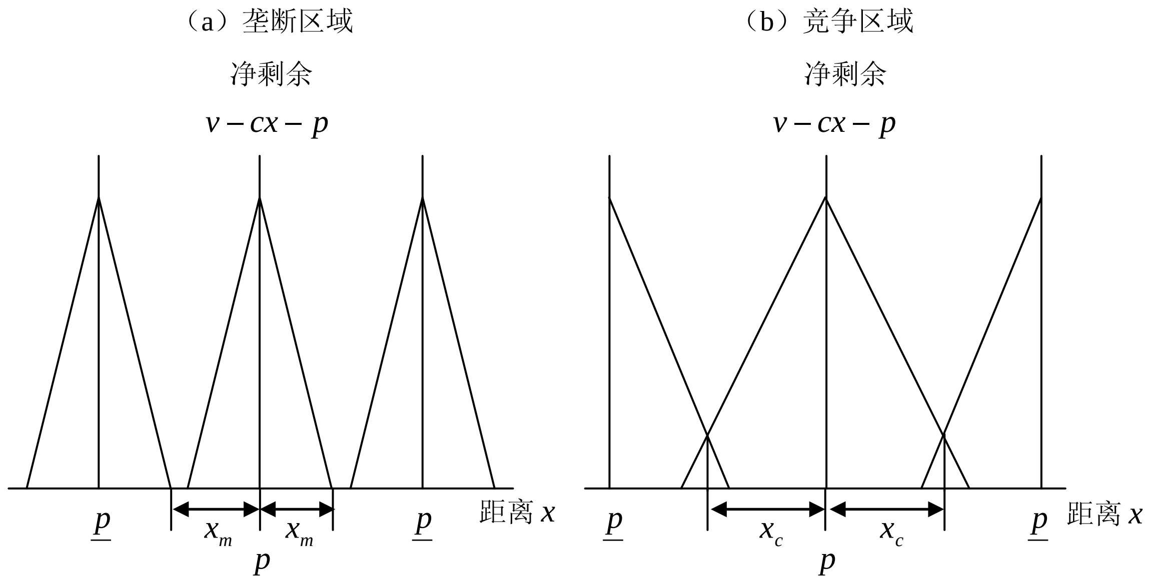


图 3.11　双市场结构

品牌吸引了在它位置两侧距离不大于 x_m 的所有消费者，或者说一段长为 $2x_m$ 的圆弧上的所有的顾客。如果圆周周围均匀地分布着 L 位消费者，品牌面临的垄断需求 q_m 为 $2x_m$L，或者根据方程（3.60）来替代 x_m，可得：

$$q_m = \frac{2L}{c}(v - p) \tag{3.61}$$

如方程（3.61）所示，若价格上涨 1 元，厂商的垄断产量需求将减少 $2L/c$。如果厂商将它的价格定得与最偏好这一产品的顾客的保留价格 v 一样高，产品销售将降至零。

④竞争区域

如果厂商较多，则它们距离较近且为同一顾客群展开竞争。于是每家厂商在确定自身价格时将其竞争对手的价格考虑进去。当厂商互相竞争时，一家厂商不能吸引所有偏爱冰淇淋胜过蛋糕的顾客；一些顾客转向了离他最近的两家对手。这些位于每两个品牌潜在市场中的顾客购买能提供最大净剩余的品牌。

与典型品牌最邻近的两家竞争者与典型品牌的距离为 $\frac{1}{n}$，索价为 $\underline{p}$。若典型品牌定价为 p，它的售出量将为多少？它吸引 x_c 距离内的所有消费者，在距品牌为 x_c 的点上消费者从该品牌获得的效用等于他们从它最邻近的竞争对手之一处获得的效用：

$$v - cx_c - p = v - c\left(\frac{1}{n} - x_c\right) - \underline{p} \tag{3.62}$$

方程（3.62）左端是从典型品牌中获得的净效用，右端是从另一品牌中获得的净效用（因为距离典型品牌为 x_c 的消费者距离其竞争对手品牌为 $\frac{1}{n} - x_c$）。图 3.11（b）

显示了竞争区域的界限 x_c 是如何由使方程（3.62）成立的那个点决定的。在这一点上，消费者并不在意购买哪种产品。当两个竞争对手品牌的净剩余相交时，消费者购买任一品牌。

解方程（3.62）得 x_c，并注意到竞争厂商的数量需求为 $q_c = 2x_c L$，竞争需求方程为：

$$q_c = \frac{L}{c}\left(\frac{c}{n} + \underline{p} - p\right) \tag{3.63}$$

因此价格每上涨 1 元（假定 $\underline{p}$ 为常数），竞争产量需求下降 $\frac{L}{c}$。即竞争需求曲线的斜率只是垄断需求曲线的一半。

⑤圆周模型的均衡类型

在高价下，厂商的需求区域不会重叠。每一家厂商都是一个地区性垄断者。当价格下降导致有更多的消费者对冰淇淋产生兴趣时，需求区域会重叠，于是，厂商间的竞争就开始了。垄断与竞争的需求区域见图 3.12。

在高于 $\underline{p}_m$ 的价格下，需求区域是垄断性质的：某一品牌的顾客不考虑购买其它任何品牌。在低于 $\underline{p}_m$ 的价格下，这一品牌与它最邻近的品牌展开竞争[29]。

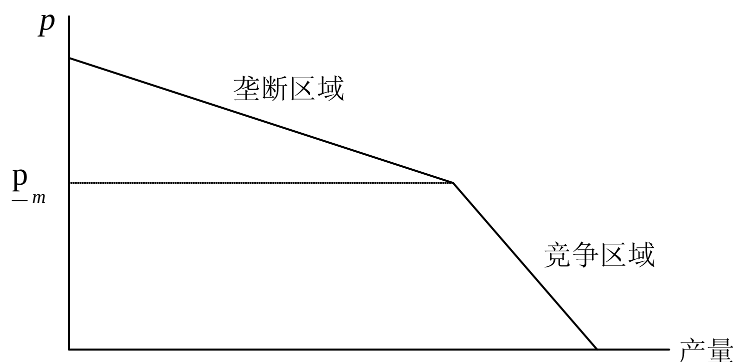


图 3.12　塞洛普圆周模型需求

塞洛普指出：当厂商的边际成本与固定成本为常数时，存在着对称的纳什均衡。行业达到这一均衡时，没有一家厂商愿意改动他的价格，也没有新增的厂商愿意进入行业。这就是说，在均衡中，所有的厂商索取相同的价格且相互间的距离为 $\frac{1}{n}$。假设允许自由进入并且厂商可无成本地重新定位以使相互间距离相等，那么，在垄断竞争均衡中，又一家厂商的进入将使所有厂商的利润为负。

⑥圆周模型成本与福利的变化

塞洛普指出，在垄断区域中，随着固定成本上升，厂商或品牌会减少，于是价格上涨，均衡多样化降低。在这一区域中，随着固定不变的边际成本上涨，价格也上涨同样的幅度（所有的成本上涨被转嫁给了消费者），但均衡多样化保持不变。

图 3.12 中价格 $\underline{p}_m$ 下的需求曲线的拗折处，无论是固定成本还是边际成本的上涨都会减少厂商数目（多样化），但有悖常理的是价格也降低了（在图 3.12 中，价格需求曲线向右下方拗折）。因此，如果经济处于这样的一个点上，提高厂商成本的税收将降低价格并降低多样化。但是塞洛普证明，即使忽略税收收入，福利也会上升。

塞洛普指出最佳最优的多样化既低于垄断均衡的多样化又低于竞争均衡的多样化。在较少的品牌下，固定成本的节约超出了高价引起的损失。因此，在圆周模型的垄断竞争均衡中毫无疑问地存在着太多的品牌。

塞洛普指出，假定政府唯一的管制政策是控制进入，则次佳最优是市场均衡或完全垄断。意即最优进入政策或是自由进入，或是限制进入从而使每一品牌都占有一个完全垄断的市场。

注释：

[1]倘若我们对剔除劣战略过程中的剔除方式不作如此限制，则定理不成立。我们在这里给出一个反例，其中对一个战略空间中劣战略的剔除并未在一次剔除过程中将所有劣战略都剔除掉。

假设博弈中有两个局中人 1 和 2，他们各自的战略空间都为 $S_1 = S_2 = \left\{ \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \cdots, \frac{1}{2^k}, \cdots \middle| k = 1,2,\cdots \right\}$，各自的支付函数为

$u_1(s_1, s_2) = k$，$s_1 = \frac{1}{2^k}$，$k = 1,2,\cdots$

$u_2(s_1, s_2) = k$，$s_2 = \frac{1}{2^k}$，$k = 1,2,\cdots$

显然，当 $s_1 = \frac{1}{2}, \frac{1}{4}, \cdots, \frac{1}{2^k}, \cdots, k = 1,2,\cdots$，$s_1$ 都是 1 的劣战略；

且当 $s_2 = \frac{1}{2}, \frac{1}{4}, \cdots, \frac{1}{2^k}, \cdots, k = 1,2,\cdots$，$s_2$ 都是 2 的劣战略。

现在，我们按这样一种劣战略剔除顺序进行劣战略剔除，即第一次剔除 1 的劣战略 $s_1 = \frac{1}{4}$，第二次剔除 2 的劣战略 $s_2 = \frac{1}{4}$，第 3 次剔除 1 的劣战略 $s_1 = \frac{1}{8}$，第 4 次剔除 2 的劣战略 $s_2 = \frac{1}{8}$，……。显然，按此顺序进行劣战略剔除，得到的重复剔除占优均衡为 $\left\{ s_1 = \frac{1}{2}, s_2 = \frac{1}{2} \right\}$。但假如我们将劣战略剔除顺序改为第一次剔除 1 的劣战略 $s_1 = \frac{1}{2}$，第二次剔除 2 的劣战略 $s_2 = \frac{1}{2}$，第三次剔除 1 的劣战略 $s_1 = \frac{1}{4}$，第四次剔除 2 的劣战略 $s_2 = \frac{1}{4}$，……。此时得到的重复剔除占优均衡是不存在的，博弈不是重复剔除占优可解的，因为局中人战略空间中的元素全部都被剔除掉了。又若我们将剔除劣战略顺序改为第一次剔除 1 的劣战略 $s_1 = \frac{1}{8}$，第二次剔除 2 的劣战略 $s_2 = \frac{1}{8}$，第三次剔除 1 的劣战略 $s_1 = \frac{1}{16}$，……。最后，我们再分别剔除 1 的劣战略 $s_1 = \frac{1}{2}$ 和 2 的劣战略 $s_2 = \frac{1}{2}$。此时又得到的重复剔除占优均衡为 $\left\{ s_1 = \frac{1}{4}, s_2 = \frac{1}{4} \right\}$。显然，这个博弈应该没有均衡，不是重复剔除占优可解的。

[2]显然，如果每一个局中人的战略空间是有限的，即战略空间中的元素个数有限，我们就不需对剔除方式作定理证明中给出的特别规定，因为在证明 $\widetilde{S}^* \supset S^*$ 过程中没有用到这一规定，仅在证明 $\widetilde{S}^* \subset S^*$ 过程中才使用该规定。但倘若 S_i 是有限空间，则 $H(m)$ 仍能定义，定理结论可类似证明。在注释[1]中给出的反例中，S_i 恰好是无限空间，所以成为定理的反例。另外，我们在定理证明中还潜在地假定了局中人个数 n 是有限的，这

一假设将贯穿全书，因为从定理具有实际意义上来看，这一假设不存在什么可挑剔的。

[3]（U, L）和（U, R）都是纳什均衡。另外，我们在这里进行的剔除方式并未按定理 3.1 中规定的方式操作，即这里的剔除并未在对 A 的非严格劣战略进行剔除时一齐将 H 和 D 剔除，但这里的例子实际上是有限博弈，即使按例子中给出的顺序剔除，定理 3.1 仍成立，见本章注释[2]。

[4]我们在这里对故事的描述在文字和情节上与《三国演义》略有不同，但故事演进的逻辑与《三国演义》中的描写是相同的（见罗贯中，1984.9）。这里的文学处理是为了使故事更加生动和适于使用博弈论语言进行分析。

[5]在《三国演义》中，关羽在华容道上捉住了曹操，但感于曹操在过去曾厚待关羽，仁义忠勇的关云长终于放过曹操一马，让其平安回到大本营许都。当时，曹操曾以过去收降关羽后对其倍加呵护的旧事感化关羽放行，但关羽以助曹操解白马之围并亲斩袁绍帐下名将颜良、文丑之功绩说明已还了曹操的旧情。聪明的曹操又以关羽闯五关斩六将时未派兵围堵且在关键时刻传令将关羽放行为由，提醒关羽仍还欠着他一笔旧情呢！关羽闻言沉默不语，片刻后挥刀令士兵让开了一条路。在这种“公平交易”下，曹操才保住了一条命。

[6]尽管如此，纳什在萌生出纳什均衡概念时却颇费周折。在不满于冯·诺依曼等人将博弈论研究的重点仅限于零和博弈的情绪中，纳什经过长期的苦思冥想后终于在 1949 年 10 月悟出了“纳什均衡”的概念。这种概念被称为是对人类行为的非凡洞察（见《普林斯顿的幽灵——纳什传》第十章，Sylvia Nasar 著，王尔山译，上海科技教育出版社，2000.10）。

[7]我们在这里潜在地假定稳定的均衡是存在的且唯一。满足这个要求的条件是：利润函数严格凹（$\pi_i'' < 0,\ \ i = 1,2$）且交叉偏导数为负$\left(\dfrac{\partial^2 \pi_i}{\partial q_i \partial q_j} < 0, i \neq j, i,\ j = 1,2\right)$；这两个条件意味着反应函数 $R_i(i = 1,2)$ 是斜率为负的连续函数；此外，还要求两曲线仅交叉一次且在交叉点 R_1 比 R_2 更陡。满足这些条件的古诺模型是重复剔除占优可解的。如果这些条件不满足，我们就无法用重复剔除的办法找到均衡解。此外，如果存在 3 个以上的寡头企业，重复剔除也无法给出均衡解。

[8]见冯·诺依曼与莫根斯坦（1944）。

[9]在我们引入混合战略之前，到目前为止我们所提到的战略都是纯战略。

[10]有人指出，根本就不存在所谓的“集体理性”，只有个人理性才是基本的假设概念。但若我们将“集体理性”理解为帕累托最优的战略组合，则这里的讨论在逻辑上不存在问题。

[11]物理学家尼尔斯·玻尔在大学时代曾被一个被称为“饿死的驴”的哲学问题所困惑：假定有一头驴，在它的面前放有两堆“完全相同”的草料。假设这头驴是“理性

的”，它的所有决定都依照“充分的”逻辑依据来作出。现在，这两堆草既然是“完全一样”的，因而它就没有“充分的”理由作出先吃哪一堆草的决定，于是，驴在犹豫不决中被活活饿死了。也许正是玻尔在青年时代的哲学思考使他在后来的量子力学理论中萌发出“上帝在掷骰子”的随机性思想。在“饿死的驴”的解决方案中，玻尔认为驴子只有采取随机性决策方法才能避免被饿死，譬如，驴子可以通过象足球比赛开球时裁判通过掷硬币决定哪一方先开球的方式来决定先吃哪一堆草，但这样一来，驴子就不是按故事开始所设定的那样机械行事了，而是有了某种“自主性”，因为它没有按照某种给定的逻辑推演来作出决定。在玻尔后来提出的量子论哲学中，他认为微观粒子在物理量上有一定的任意选择性。譬如，在著名的粒子通过窄孔的衍射实验上，通过窄孔的粒子可以打在衍射屏上的任意位置上。在海森堡提出的“测不准原理”中，粒子的所有（共轭）物理量（如位置和速度）是不能同时精确测定的。尽管爱因斯坦至死都极力反对量子论的哲学观点，认为“上帝不会掷骰子”，但事实似乎已证明以玻尔为首的哥本哈根学派最终是正确的 ——上帝似乎是在通过掷骰子来安排宇宙的秩序。

在弱纳什均衡中，我们碰到与“饿死的驴”相似的问题：倘若局中人有两个无差异的最优战略，在我们假定的“理性人”条件下，局中人没有充分的理由决定是选择其中某一战略呢还是选择另一战略。当这个局中人正好是玻尔的“驴子”时，问题就完全一样了。所以，为了避免出现“饿死的驴”那样的结局，我们将目光集中于强纳什均衡而不去管弱纳什均衡，这也许正是善于通过假设来回避问题的经济学家们惯用的伎俩，也就是说在玻尔的驴这个故事中假定两堆草是不一样的，这样就将问题取消了。当然，我们也可以假定局中人在两个无差异的最优战略之间随机地选择，这就是后面将要引入的“混合战略博弈”，但这样一来，理性人假设又重新面临挑战，因为这种随机性决策似乎又与博弈论（抑或经济学）关于理性人的假定相冲突。我们将在后面介绍“混合战略博弈”时重新回到这种讨论上来。不过，这里可以事先提醒读者，关于这方面的问题讨论已进入哲学范畴，目前仍未有一致的结果。

[12]这里显然有 $H=P$，$2^H-H=2^P-P$，故 $a=b$ 可能实现。

[13]类似现象还有，一些人总是“小错不断，大错不犯”。

[14]有趣的是，公共品供给问题中既有囚徒困境现象，又有智猪博弈特点，还可能是斗鸡博弈，依具体的支付函数而定。

[15]在这里，与式（3.7）有一点不同的是我们用 p_j 表示局中人 j 的纯战略空间中的某一个纯战略，而 p_j 正是其所表示的纯战略所对应的指标。

[16]我们用 S_j 中元素对应指标 p_j 代表元素本身。

[17]参见 3.2.12 小节中关于“奇数定理”的说明。

[18]当局中人认为不确定情况下的期望支付与确定性情况下的同一支付给他带来同样的效用时，称局中人是风险中性的。当前者带来的效用小于后者带来的效用时，称局

中人是风险规避的；相反，当前者带来的效用大于后者带来的效用时，称局中人是风险偏好的。

[19]如张维迎（1996）。

[20]纳什存在性定理保证该博弈至少存在一个纳什均衡，但该博弈的纳什均衡是混合战略博弈均衡。

[21]运用公共品私人供给不足原理似乎还可以对一些其它现象作出有趣的解释，譬如，城市街道两边的私人房屋的外观欠美观和肮脏，大街上好看的女孩子们太少等等。

[22]Bertrand 寡头模型是 Bertrand 于 1883 年提出。

[23]在这里给出的线性需求函数情况下，$P_{1\max}$ 和 $P_{2\max}$ 都是确定的有限数，令 $q_1=0$，$q_2=0$，解得 $P_{1\max}=\dfrac{a_1b_1+d_1a_2}{b_1^2-d_1d_2}$，$P_{2\max}=\dfrac{a_2b_2+d_2a_1}{b_2^2-d_1d_2}$。

[24]这里未对产品完全可替代情形进行计算，但在这种情形厂商面临的需求函数的表达较为复杂，见 Bertrand（1883）。

[25]与价格竞争类似，这个模型研究的重点尽管是地理位置，但通过将“位置”理解为刻画产品差异的特征可以用来研究厂商的差异化竞争即垄断竞争。

[26]假定两个商店不能位于相同的地理位置。

[27]与例 3.30 一样，假定每个消费者都购买一个单位产品。

[28]存在一个混合战略均衡，在这个均衡中，每家厂商按照一定概率选择它的行动。但是，作为产品竞争来说，混合战略不能解释观察到的现象，因为一旦决策作出并实现后，局中人发现在给定对方的选择下自己的选择不是最优的，从而带来一系列无止境的调整，而这种情况并不是真实出现的。

[29]如方程（3.62）所示，在价格低于 $\underline{p}-\dfrac{c}{n}$ 的情形下，所有在一个给定厂商和其相邻厂商之间的消费者会倾向于从低价厂商处购买，即使他们不怎么喜欢这个品牌。考虑一下在相邻厂商位置上的顾客：他与低价厂商的距离为 $\dfrac{1}{n}$。这个消费者通过消费低价品牌而不是他所喜欢的品牌而损失了 $\dfrac{c}{n}$ 的的效用。但是价格所带来的节省要大于那个损失。这类行为是极具进攻性的（或者像塞洛普所谓的“超竞争性”），因为低价厂商攫取了所有新的相邻厂商的消费者。为简明起见，在图 3.12 中没有显示这个地区。

第四章　纳什均衡的存在性与多重性

对于数学家来说，一个数学概念的存在性与唯一性是特别需要加以关注的。这是因为，从形式逻辑角度看，如果某个事物并不存在，那么关于这个杜撰中的事物所给出的任何陈述或判断都可认为是正确的或错误的，因为对于不存在的事物来说，任何关于它的陈述或判断都不可能加以证伪。所以，倘若某个概念所对应的事物并不存在。那么，关于这个概念所给出的研究结论都必然不存在被证伪的可能。因而根据波普尔的证伪主义观点，这样的研究不具备科学上的意义。所以，我们在对任何新提出来的数学概念加以系统研究之前，首先需要弄清楚所研究的对象事物是否存在。

有许多被称为伪科学的东西，它们之所以被人们认为是"伪科学"的原因就是它们大肆谈论的东西并不存在或并未被证实其存在性。譬如，所谓的特异功能或"超灵学"并未得到证实，而 UFO 研究迷们至今也未能拿出一件存在球外生命的证据，所以，特异功能学或"超灵学"或"不明飞行物学"实际上都可被归入伪科学。除了存在性之外，概念事物的唯一性也是数学家们所关心的问题。从纯理论的兴趣上看，数学家们更多地是从审美的角度上看待概念的唯一性，但从波普尔的证伪主义哲学看，模型均衡解的唯一性关系到模型的预测功能，从而是科学理论应基本具有的特征。我们在第二章中曾指出，理论的预测功能是判别理论的科学性的准绳，而在第三章中，我们提出用纳什均衡作为模型的预测结果。按照这样的逻辑，一个自然的推论就是：模型能否具有科学意义取决于纳什均衡的唯一性。因为倘若纳什均衡不是唯一的，那么就难以根据模型对即将出现的结果加以预测，这种不确定性对于科学理论来说是不存在的。再加上前面谈到的存在性问题，我们可以这样说，模型能否具有科学意义取决于纳什均衡的存在性和唯一性，因为这正是科学理论所具有的基本性质。

博弈论目前发展的情况是这样的：已经证明在非常一般的情况下，纳什均衡是存在的，这是一个好的结果；但是，在许多情形，模型的纳什均衡解不是唯一的，这被称为纳什均衡的多重性问题。

纳什在 1950 年代证明了纳什均衡的存在性定理，为非合作博弈打下了重要基础。纳什的工作不仅解决了存在性问题，而且还为其后的博弈论研究提供了一整套方法论工具，即运用不动点定理(fixed point theorem)这一强有力的数学工具进行博弈论数学分析，这对后来的博弈论甚至数理经济学的发展产生了很大的影响。纳什均衡的多重性问题至今仍是困扰博弈论学者的一个主要问题。为了攻克这一问题，博弈论专家已经做出了许多贡献，如聚点均衡、相关均衡，子博弈精炼纳什均衡，颤抖手均衡，序贯均衡等概念的提出。但不幸的是，这类努力还未使得多重均衡问题完全得到解决，许多博弈论专家正在这一领域进行着不懈的工作。

本章将给出纳什均衡的存在性定理和讨论存在多重均衡情况下的均衡选择问题。

4.1 纳什均衡的存在性定理

自从纳什(1950)首先给出存在性定理及其证明之后，许多学者又相继提出了不同表述下的存在性定理和不同的证明方法。这里，我们介绍 Myerson(1991)给出的存在性定理和证明。

4.1.1 纳什均衡与不动点定理

所有的存在性定理证明都采用了不动点定理，这是因为，纳什均衡的概念在数学上就是一个不动点的概念。在给出存在性定理及其证明之前，我们先来说明不动点的概念和给出不动点定理。

什么是“不动点”呢？考虑一个方程 $f(x)=x$，其中 x 为方程的解。我们将 $f(\cdot)$ 视为一种“变换”，即 $f(\cdot)$ 是将 x 对应为 $y=f(x)$ 的变换，其中 x 和 y 分别是属于集合 X 和 Y 的两个元素，$x\in \mathrm{X}$，$y\in Y$。如果 $X=Y$，则方程 $f(x)=x$ 的几何意义就是：变换 $f(\cdot)$ 将 x 变为自己，即 x 在 $f(\cdot)$ 变换下是不变的，故称 $f(x)=x$ 的解为变换 $f(\cdot)$ 的不动点。

一般地，我们可以将所有的方程都写为如下形式：

$$y(x)=0 \tag{4.1}$$

在式(4.1)两端加上一个 x，则变为 $y(x)+x=x$。

令 $f(x)=y(x)+x$ 则有

$$f(x)=x$$

所以，一般地，方程求解的问题本质上是寻找变换的不动点问题。

对于这样一种非常一般地的问题，数学家们感到十分高兴的是居然在不太严格的条件下式(4.1)存在解，即不动点是较为广泛地存在的。

譬如，图 4.1 表明不动点是曲线 $f(\cdot)$ 与 45^{o} 线的交点。当函数 $f(x)$ 定义在 $x\in[0,1]$ 区间上且因变量 $y=f(x)$ 的值域也为 $[0,1]$ 区间时，如果 $f(x)$ 是连续的，则必然存在不动点。

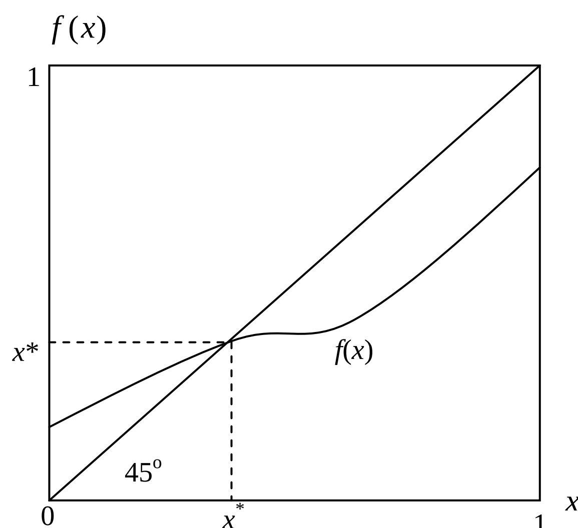


图 4.1 [0,1]区间上的自变换函数的不动点

那么，这种现象到底具有多大的一般性意义呢？

数学家 Brouwer 在很久以前就注意到这一现象，他得出了如下的一般性定理，即著名的 Brouwer 不动点定理。

定理 4.1(Brouwer……)

设 $f(x)$ 是定义在集合 X 上的实函数，且 $f(x)\in X$，$\forall x\in X$。

如果 $f(x)$ 是连续的，X 为一非空的有界凸闭集，则至少存在一个 $x^*\in X$ 使 $f(x^*)=x^*$。即 $f(x)$ 至少存在一个不动点[1]。

有意思的是，Brouwer 不动点定理存在很强的几何直观[2]，但其数学证明却十分艰深，需要动用代数拓扑这类就是职业数学家也感到望而生畏的超级抽象数学工具[3]。

在此，我们不给出 Brouwer 不动点定理的证明。

直接用来证明纳什存在性定理的不动点定理还不是 Brouwer 不动点定理，而是角谷静夫(Kakutani)不动点定理，而后者的证明只是前者的一个相对简单的运用。

我们所以要引用角谷静夫不动点定理，是因为在纳什均衡存在性证明中所遇到的反应函数一般是多个因变量函数，即所谓对应(correspondence)，而角谷静夫不动点定理正好描述的是对应的一种性质。角谷静夫不动点定理是 Brouwer 不动点定理的推广，但其自身的证明要用到 Brouwer 不动点定理。我们在这里不打算给出这两个不动点定理的证明，因为这类证明只是一种纯数学过程，但我们将给出纳什存在性定理的一种证明，因为了解存在性定理的证明过程有助于我们更好地理解纳什均衡。

为了解读角谷静夫不动点定理，我们先来准备一下一些有关的数学概念。

对于任一有限集 M，我们用 R^M 表示形如 $x=(x_m)_{m\in M}$ 的所有向量组成的集合，其中对 M 中每一个 m，第 m 个分量 x_m 是实数域 R 的一个元素。为方便计，我们也可将 R^M 等价地理解为 M 到 R 上的所有函数组成的集合，这时 R^M 中 x 的 m 分量 x_m 也可被记为 $x(m)$。

令 S 是 R^M 中的一个子集，我们有如下定义：

定义 4.1　S 是凸的(Convex)当且仅当对任意的 $x\in R^M, y\in R^M$ 及满足 $0\le\lambda\le 1$ 的 λ，只要 $x\in S$ 和 $y\in S$，则有

$$\lambda x+(1-\lambda)y\in S$$

这里，$x=(x_m)_{m\in M}, y=(y_m)_{m\in M}, \lambda x+(1-\lambda)y=(\lambda x_m+(1-\lambda)y_m), m\in M$

定义 4.2，S 是闭的(Closed)当且仅当对每个收敛的序列 $\{x(j)\}_{j=1}^{\infty}$，如果对每个 j 都有 $x(j)\in S$，则有

$$\lim_{j\to\infty} x(j)\in S$$

定义 4.3，R^M 中的子集 S 是开的(open)当且仅当它的补集 R^M/S 是闭的。

定义 4.4，S 是有界的(bounded)当且仅当存在某个正数 K 使得对 S 中的每个元素 x 都有

$$\sum_{m\in M}\left|x_m\right|\le K$$

定义 4.5，一个点到集合的“对应”(correspondence) $G:X\to Y$ 是任何一个规定了对 X 中的每个点 x， $G(x)$是与 x 相对应的 Y 中的一个子集。

如果 X 和 Y 都是度量空间，则 X 和 Y 上的收敛和极限概念已经定义，这时有：

定义 4.6 ，一个对应 G:X→Y 是上半连续的(upper—hemicontinuous)，当且仅当对每个序列 $\{x(j),y(j)\}_{j=1}^{\infty}$，如果对于每个 j 有 $x(j)\in X$ 和 $y(j)\in G(x(j))$，而且序列 $\{x(j)\}_{j=1}^{\infty}$ 收敛于某个点 $\bar{x}\in X$ ，又序列 $\{y(j)\}_{j=1}^{\infty}$ 收敛于某个点 $\bar{y}\in Y$，则有

$\bar{y}\in G(\bar{x})$

定理 4.2，对应 $G:X\to Y$ 是上半连续的当且仅当集合 $\{(x,y)|x\in X,y\in G(x)\}$ 是集合 $X\times Y$ 中的一个闭子集。

证明：必要性。记集合 $A=\{(x,y)|x\in X,y\in G(x)\}\subset X\times Y$.

设 $Z_j=(x(j),y(j))$ 为 A 中一收敛序列，其中 $x(j)\in X$,

$y(j)\in g(X(j)),\ j=1,\cdots,\infty$

由上半连续性知 $\lim_{j\to\infty}y(j)\in G\left(\lim_{j\to\infty}x(j)\right)$

显然有 $\lim_{j\to\infty}x(j)\in X$

故 $\lim_{j>\infty}Zj\in A$，所以 A 为 $X\times Y$ 中一闭子集。

充分性。假设 A 为 $X\times Y$ 上的一个闭子集。

如果序列 $\{x(j),y(j)\}_{j=1}^{\infty}$ 中每个 $x(j)$ 和 $y(j)$ 都有

$x(j)\in X$ ，

$y(j)\in G(x(j))$

且 $\{x(j)\}_{j=1}^{\infty}$ 收敛于 $\bar{x}$ 和 $\{y(j)\}_{j=1}^{\infty}$ 收敛于 $\bar{y}$，则 $Z_j=(x(j),y(j))$ 收敛于 $(\bar{x},\bar{y})$。

由 A 的闭性知 $(\bar{x},\bar{y})\in A$，即 $\bar{y}\in G(\bar{x})$

故 G 为上半连续。

证毕！

上半连续性是我们熟知的连续函数概念的一种推广，而函数的连续性比上半连续性要强一些，于是有

定理 4.3，如果 $y:X\to Y$ 是一个从 X 到 Y 的连续函数，且对 X 中的每一个 X 都有 $G(x)=\{y(x)\}$，那么 $G:X\to Y$ 是一个点到集的上半连续对应。

证明：

设序列 $\{x(j),y(j)\}_{j=1}^{\infty}$，且对每个 j 有 $x(j)\in X$ 和 $y(j)\in G(x(j))$， $\{x(j)\}_{j=1}^{\infty}$ 收敛于 $\bar{x}$，$\{y(j)\}_{j=1}^{\infty}$ 收敛于 $\bar{y}$ 。

由 y 的连续性知 $\overline{y} = y(\overline{x})$

故 $\overline{y} \in G(\overline{x})$

于是 G 是上半连续的。

下面，我们将不动点概念扩充到对应的情形。

定义 4.7，一个对应 F：$S \to S$ 的一个不动点是 S 中任一满足 $x \in F(x)$ 的 x。

角谷静夫得出如下被广泛应用的一个重要定理。

定理 4.4（角谷不动点定理）令 S 是一个有限维向量空间中任一非空有界闭凸子集。设 F：$S \to S$ 是任一上半连续的点到集对应，且对 S 中每个 $x, F(x)$ 都是 S 的一个非空凸子集。那么，S 中一定存在某个 $\overline{x}$ 使得 $\overline{x} \in F(\overline{x})$(Kakutani, 1941)

角谷不动点定理说的是对于有限维向量空间中任一非空有界闭凸子集上的上半连续自对应来说，在一定条件下都至少存在一个不动点。角谷不动点定理及其它的一系列相关定理的证明还可参见 Burger(1963), Franklin (1980)和 Border(1985)。数理经济学家 Scarf(1973)曾通过一种计算不动点的算法而提供了一个构造性证明，其中不动点的存在性是由这个定理所保证的。关于角谷不动点定理的推广，可参见 Glicksberg (1952)。

4.1.2 纳什存在性定理及其证明

下面，我们来证明纳什存在性定理，该定理最早由纳什得出，这里的证明由 Myerson(1991)给出[5]。

定理 4.5 (Nash , 1950)，任何一个战略式表述的有限博弈都至少存在一个混合博弈纳什均衡。

证明：令 Γ 是任—战略式表述有限博弈，即

$\Gamma = \{S_1 \cdots S_n; u_1 \cdots, u_n\}$

显然，$\sum = \prod_{i=1}^{n} \Sigma_i$ 是一个有限维向量空间的一个非空有界闭凸子集(注意 Γ 是有限博弈，即局中人数和每个 S_i 中的元素个数都是有限数)[6]。

任给 $\sigma \in \sum$ 和任一局中人 i，令

$R_i(\sigma_{-i}) = \arg\max_{\sigma_i \in \Sigma_i} V_i(\sigma_i, \sigma_{-i})$

即 $R_i(\sigma_{-i})$ 是局中人 i 在 $\sum_i$ 中对其余局中人独立混合战略组合 σ_{-i} 的最优反应混合战略。

根据定理 3.2，$R_i(\sigma_{-i})$ 是 S_i 上所有的概率分布 σ_i 组成的集，且使得对每一个满足 $S_i \overline{\in} \arg\max_{s_i \in S_i} V_i(s_i, \sigma_{-i})$ 的 S_i 有 $\sigma_i(s_i) = 0$，由定理 3.2 的证明过程知道，

$V_i(\sigma_i, \sigma_{-i}) = \sum_k \sigma_{ik} V_i(s_{ik}, \sigma_{-i})$

任给 $\sigma_i \in R_i(\sigma_{-i}), \sigma_i^{'} \in R_i(\sigma_{-i}), \lambda \in [0,1]$, 令

$\sigma_i^{''} = \lambda\sigma_i + (1+\lambda)\sigma_i$

显然 $\sigma_i^{"} \in \sum_i$，

$V_i\left(\sigma_i^{"}, \sigma_{-i}\right) = \sum_k \sigma_{ik}^{"} V_i\left(S_{ik}, \sigma_{-i}\right)$

$= \lambda \sum_k \sigma_{ik} V_i\left(s_{ik}, \sigma_{-i}\right) + (1-\lambda) \sum_k \sigma_{ik}^{'} V_i\left(s_{ik}, \sigma_{-i}\right)$

$\geq \lambda v_i\left(\tilde{\sigma}_i, \sigma_{-i}\right) + (1-\lambda) V_i\left(\tilde{\sigma}_i, \sigma_{-i}\right)$

$= V_i\left(\tilde{\sigma}_i, \sigma_{-i}\right), \forall \tilde{\sigma}_i \in R_i(\sigma_{-i})$

故 $\sigma_i^{"} \in R_i\left(\sigma_{-i}\right)$，所以 $R_i\left(\sigma_{-i}\right)$ 是凸的。

根据 $V_i\left(\sigma_i, \sigma_{-i}\right) = \sum_R \sigma_{ik} V_i\left(s_{ik}, \sigma_{-i}\right)$，因为 S_i 是有限集，故存在某个 k 使 $V_i\left(s_{ik}, \sigma_{-i}\right) = \max_l \left[V_i\left(s_{il}, \sigma_{-i}\right)\right]$

即 argmax $V_i\left(s_i, \sigma_{-i}\right)$ 是非空的。令 $\sigma_{ik} = 1, \sigma_{il} = 0, l \neq k$，则

$V_i\left(\sigma_i, \sigma_{-i}\right) = \max_{\sigma_i \in \sum_i} V_i\left(\sigma_i, \sigma_{-i}\right)$

即 $\sigma_i \in R_i\left(\sigma_{-i}\right)$

故 $R_i\left(\sigma_{-i}\right)$ 非空。

下面构造对应 R，它将 $\sum$ 中的点映射于 $\sum$ 中的子集，满足：

$R(\sigma) = \coprod_{i=1}^{n} R_i\left(\sigma_{-i}\right), \sigma \in \sum$

由于对每一个 $i = 1, \cdots, n$，$R_i\left(\sigma_{-i}\right)$ 都是非空凸集，显然 $R(\sigma)$ 也是非空凸集。下面我们来证明 R 是上半连续的。

假设 $\left\{\sigma^k\right\}_{k=1}^{\infty}$ 和 $\left\{\tau^k\right\}_{k=1}^{\infty}$ 都是收敛序列

$\sigma^k \in \sum$，$\tau^k \in R\left(\sigma^k\right)$

$k = 1, 2, \cdots$

且 $\overline{\sigma} = \lim_{k \to \infty} \sigma^k, \overline{\tau} = \lim_{k \to \infty} \tau^k$

为了证明 R 是上半连续的，我们将需要证明 $\tilde{\tau} \in R(\overline{\sigma})$。

因为有：

$V_i\left(\tau_i^k, \sigma_{-i}^k\right) \geq V_i\left(\sigma_i, \sigma_{-i}^k\right), \forall \sigma_i \in \sum_i$

$k = 1, 2, \cdots$

显然期望效用函数 V_i 是 $\sum$ 上的连续函数，故有

$V_i\left(\overline{\tau}_i, \overline{\sigma}_{-i}\right) \geq V_i\left(\sigma_i, \overline{\sigma}_{-i}\right), \forall \sigma_i \in \sum_i$

因此，对于每一个 i 有 $\overline{\tau}_i \in R_i\left(\overline{\sigma}_{-i}\right)$, 故

$\overline{\tau} \in R\left(\overline{\sigma}\right)$。

所以 R 是 $\sum$ 到自身上的一个上半连续对应。

根据角谷不动点定理，存在 $\sum$ 中的某个混合战略组合 σ 使 $\sigma \in R(\sigma)$，即对于每一个 i 有 $\sigma_i \in R_i\left(\sigma_{-i}\right)$，因此 σ 就是 Γ 的一个（混合）纳什均衡。

证毕！

4.1.3 其它的纳什均衡存在性定理

在纳什存在性定理中，我们只谈及到包括混合战略均衡在内的纳什均衡存在性问题，除此之外，我们自然会对纯战略纳什均衡的存在性感到特别的兴趣。另外，许多博弈不一定是有限博弈，一些常见的博弈的纯战略空间通常都是无限集。在纳什定理之后，其他研究者还得到许多进一步的结果，这些结果中与上述问题相关的有如下几个定理。

定理 4.6(Debreu, 1952; clicksberg, 1952, Fan, 1952) 在 n 人战略式表述博弈 $G=\{S,\cdots,S_n;u_1,\cdots,u_n\}$ 中，如果纯战略空间 S_i 是欧氏空间上的非空有界闭凸子集，支付函数 u_i 是连续的且对 S_i 是拟凹的 $(i=1,\cdots n)$，则 G 存在一个纯战略纳什均衡。

一般地，当函数 $f(x)$ 满足下述性质时，我们称其为凹的：

$$f(\lambda x_1+(1-\lambda)x_2)\geq \lambda f(x_1)+(1-\lambda)f(x_2),\lambda\in[0,1]x_1,x_2\in R^n$$

如果当 $\lambda\in(0,1)$ 时上面的不等式严格成立，则称 $f(x)$ 为严格凹的。一个函数 $f(x)$ 是凸的当且仅定函数- $f(x)$ 是凹的； $f(x)$ 为严格凸函数当且仅当- $f(x)$ 为严格凹函数。

拟凹函数是凹函数概念的一种推广，它包括了凹函数在内的一大类函数，而这类函数在经济学中有着广泛应用，关于拟凹函数的定义如下：

定义 4.8，函数 $f(x)$ 定义在 R^n 中的子集 D 上，当且仅当 $f(x)$ 满足如下性质时，$f(x)$ 是拟凹的：

$$f(\lambda x_1+(1-\lambda)x_2)\geq \min(f(x_1),f(x_2))\qquad \lambda\in[0,1]$$

显然，凹函数是拟凹的，但反过来并不成立，即拟凹函数不一定是凹函数。在图 3.2 中，函数 $f(x)$ 是拟凹的，但不是凹的。

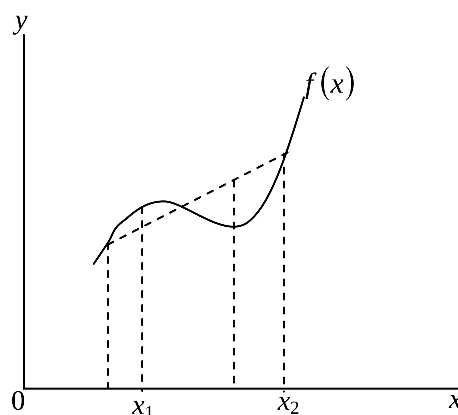


图 4.2 不是凹函数的拟凹函数

在定理 4.6 中，与定理 4.5 相比，我们增强了对支付函数 u_i 性质的假设，于是获得更进一步的结论，即保证了存在的纳什均衡还是纯战略博弈纳什均衡。在有限博弈场合，即使纯战略空间可能是非凸的，支付函数也可能是非连续的，但混合战略空间是欧氏空间上的非空有界闭凸集，期望支付函数是连续的，拟凹的。当纯战略空间本身是欧氏空

间上一个非空的，闭的，有界的凸集且支付函数在纯战略空间上是连续的，拟凹的时，就没有必要引入混合战略了。

如果放松定理 4.6 中关于支付函数的拟凹性假设，则只能保证混合战略均衡的存在性，这就是下面的定理 4.7。

定理 4.7 (Glicksberg, 1952) ，在 n 人战略式表述博弈 $G=\{S_1,\cdots,S_u;u_1\cdots u_n\}$ 中，如果纯战略空间 S_i 是欧氏空间上一个非空有界闭凸集，支付函数 u_i 是连续的，则 G 存在一个混合战略纳什均衡。

注释：

[1]这个定理的表述中隐含了 X 为一个度量空间，所谓度量空间，即在空间 X 上定义了一个“距离”函数$\|\cdot\|$，使得对任意的$x_1 \in X, x_2 \in X$都有

(i)　　$\|x_1 + x_2\| \le \|x_1\| + \|x_2\|$(三角不等式，意思是三角形的两边之和大于第三边)

(ii)　　$\|x_1\| \ge 0, \|x_2\| \ge 0$

同时还有　$\|x\| = 0$ 当且仅当 $x = 0$

当然，这种定义又要求在空间X上首先定义了一种加法“+”和“零”元素。一般地，度量空间的形式化定义为：集合 X 上的“距离”指$X \times X$到实数轴 R 上的一个函数$\delta(x, y)$，满足：对 X 中任意的x, y和 Z，有：

$$\delta(x, y) = \delta(y, x) \ge 0 \qquad (\text{对称性})$$

$$\delta(x, y) = 0 \qquad \text{当且仅当} x = y$$

$\delta(x, y) + \delta(y, z) \ge \delta(x, z)$ (三角不等式)

[2]譬如，揉面的师傅都有着这样的体验，即在面板上揉面时总有一些面粒的位置基本上不因揉面动作变化；另外，男人在梳头时总会发现某一撮头发梳不平整——它们呈竖立状伸出。

[3]某些数学家声称已找到 Brouwer 不动点定理的初等证明，但从严格的数学证明所要求的严密程度看，这类“证明”，并非真正数学意义上的证明，同时，它们还十分繁锁。譬如见……。

[4]这个定义中隐含了$\bar{x} \in X$的假设。

[5]我们这里将 Myerson(1997)中的证明作了一些形式上的修改，主要是为了适应本书的符号系统。

[6]$\sum$ 是R^M中的一个子集，而$M = \bigcup_{i=1} S_i$。

第五章　完全信息动态博弈

本章将引入博弈顺序于博弈模型之中，并对不同博弈顺序对博弈均衡的影响作出评估。正如我们在第三章中所指出，在“动态”的博弈模型中，博弈中局中人的“出招”顺序对于均衡形成来说是十分重要的。所谓“动态”博弈模型，就是指在博弈中局中人的“出招”顺序对均衡的形成有实质性影响的博弈模型。

本章将表明，影响动态博弈的均衡形成除了局中人的博弈顺序要素外，还有博弈的时间长短这一重要因素，即有限次和无限次（重复）博弈可能会有完全不同的均衡。在本章还暂不考虑信息不完全对博弈均衡的影响，因而所表达的博弈是完全信息动态博弈。

本章将要指出的重要观点是，当我们在博弈中纳入决策的顺序结构时，纳什均衡作为模型的预测是不能令人满意的，并引入子博弈精炼纳什均衡这一概念来克服这一不足。这种新概念的引入实际上是将纳什均衡概念进一步精致化，消除了一些不可信的均衡，这也是一种有限度解决纳什均衡多重性的方案（尽管还不是终极性的）。

5.1 完全信息动态博弈：概念与定义

我们在第三章已经对动态博弈概念作了一种粗略的介绍，这里，我们对完全信息动态博弈的精确定义及相关概念给出完整的说明。首先，我们通过图 5.1 给出的一个完全信息动态博弈的扩展式表示，来为即将给出的一些概念作出直观上的理解。

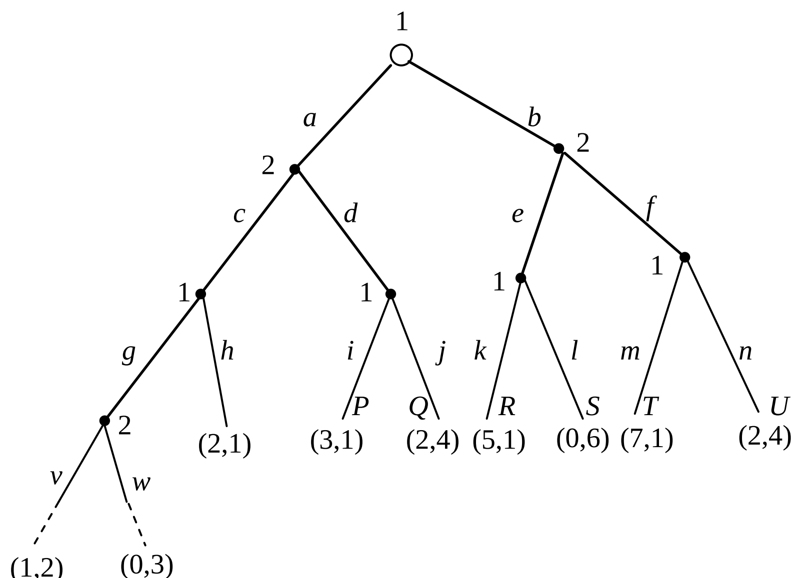


图 5.1　一个完全信息动态博弈模型

让我们看看图 5.1，在第三章中我们已给出了一些相关的概念，如决策结、枝、终点结、支付向量等等，我们称第一个行动选择对应的决策结为“初始结”。与那里给出

的例子不同的一点是，图 5.1 中存在无限个“枝”，即该博弈是一个无限次的动态博弈。在每一个决策结处，对应一个局中人，他（她）在该决策结处进行决策。即在每一个决策结处相应存在一个对应的局中人和一个对应的行动空间。对于每一个终点结，我们都可以找到一条由相互连贯的许多“枝”构成的“路径”（Path），使得该终点结是该“路径”的一个“末端”。如终点结 P 是路径“adi”的一个末端，其中 adi 指由从枝 a 经枝 d 到达枝 i 的路径。我们发现，每一个终点结对应的这种“路径”是唯一的。所以，我们可以用这种“路径”来代表终点结。但是，沿由枝 a、c、g、v、……和 a、c、g、w、……构成的“路径”却没有终点结，但图中标明，若博弈沿路径“acgv……”进行下去，局中人 1 和 2 的支付分别为“1”和“2”，若博弈沿路径“acgw……”进行下去，局中人 1 和 2 可分别获得支付“0”和“3”。从图 5.1 中还可看出，局中人 1 在博弈中总是能记住他自己过去所进行的决策，局中人 2 也是如此，同时，局中人 2 不仅能在博弈进行中总能记住自己过去所做出的决策，而且还不会忘记自己过去所知道的事，我们称这种情况为局中人具有“完美回忆”。这就是在“完全信息”基础上所进一步作出的假设。当然，“完美信息”假定还要求局中人还知道在他自己进行决策之前由别的局中人所做出的决策。

与图 5.1 的情形形成对比的是，图 5.2 给出的扩展式博弈就不具有“完美回忆”。

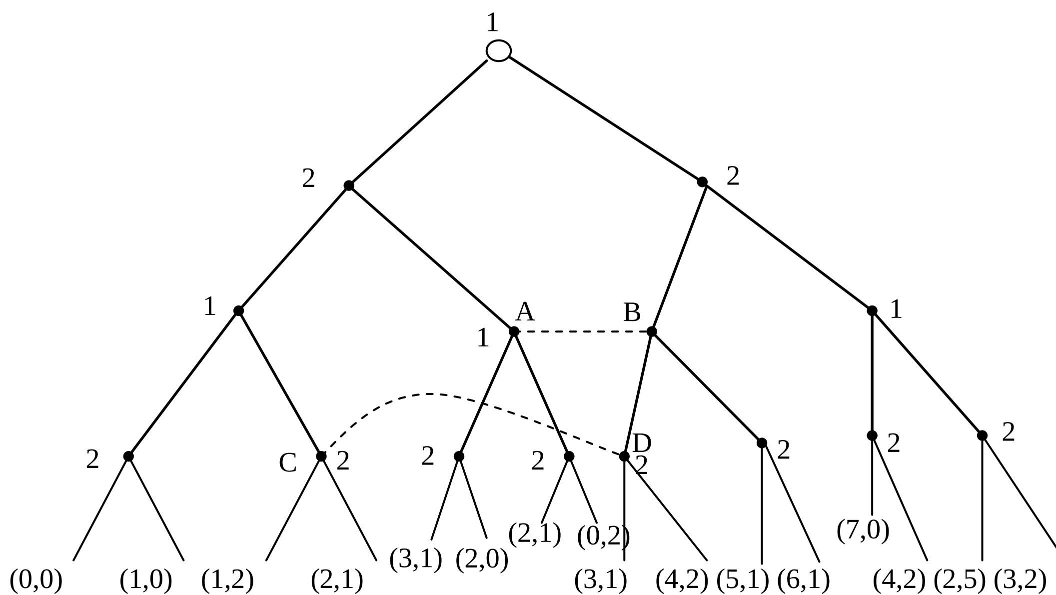


图 5.2　不具完美回忆的动态博弈

下面给出“信息集”的定义。

在图 5.2 中，局中人 1 在由决策结 A 和 B 构成的“信息集”处忘记了自己过去的决策，此时局中人 1 不知道自己到底位于 A 还是 B 处，我们称 A 和 B 构成一个“信息集”。在每一个决策结上，对应该决策结有一个唯一的局中人在此进行决策，决定其行动选择。假定每一个决策结处有一个唯一的行动空间，即由在该处进行决策的局中人进行行动选

择时所可能选择的所有行动构成的集合。如果一个信息集中含有两个或两个以上的决策结，假定其中所有决策结上对应的行动空间都是相同的，否则局中人凭此就可以知道自己在哪一个决策结上，而这与该信息集的定义相矛盾。在由决策结 C 和 D 构成的另一信息集处，局中人 2 忘记了自己知道的事，即局中人 1 在博弈开始所进行的决策。如果每一个信息集都是单结的，则称博弈为“完美信息博弈”（game of perfect information），即每一个信息集都含有唯一一个决策结。

5.1.1 完全信息动态博弈的定义

在以上直观性说明的基础上，我们可以通过运用较为抽象的术语给出完全信息动态博弈的定义。其实，按照以下的方式，我们先给出动态（或扩展式）博弈的一般性定义（不局限于完全信息的动态博弈），然后在其基础上对信息完全性作出规定就获得完全信息动态博弈的相应定义。

首先，我们借用图论中的概念对“博弈树”作出说明。

设有可数性集合 $V=\{V_i\}$，其中 $i\in I$，I 为一可数指标集。构造另一集合 $A=\{(v_i,v_j)|v_i\neq v_j,v_i,v_j\in V,i,j\in I\}$。一个未定向的“图”指由集合 V 与 A 的一个子集 $\tilde{A}$ 的一个构成[1]。称 $v_i\in V(i\in I)$ 为图的一个“顶点”（vertex）或“结点”（node），称 $(v_i,v_j)\in\tilde{A}(i,j\in I)$ 为一段（联结“顶点” v_i 和 v_j 的）弧（arc）或“枝”（branch）或“边”（edge）。一个连接结点 v_1 与 v_m 的（开的）“路径”（path）指一个结点序列 v_1，v_2，…，v_m 满足：$v_i\in V(i\in I)$ 且 $(v_1,v_2),(v_2,v_3),\cdots,(v_{m-1},v_m)$ 都在集合 $\tilde{A}$ 中[2]。

当 $v_1=v_m$（且 $m\geq 2$）时，我们就得到一个“圈”（cycle）或“闭路径”。图 5.3 中给出了两个“图”的例子。

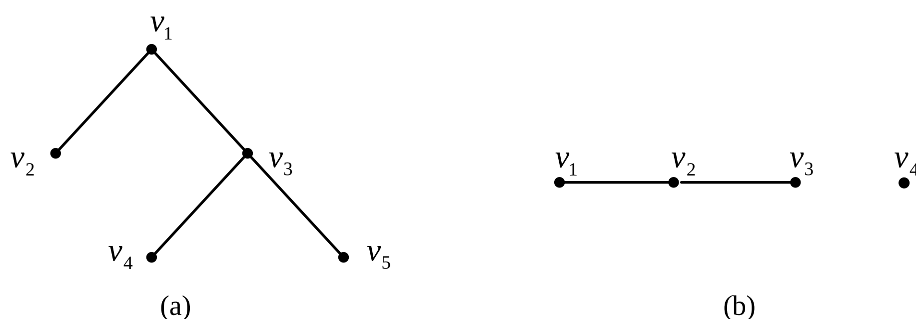


图 5.3　两个图的例子

在(a)中，$V=(v_1,v_2,v_3,v_4)$，$\tilde{A}=\{(v_1,v_2),(v_1,v_3),(v_3,v_4),(v_3,v_5)\}$；在(b)中，$V$ 与(a)中相同，而 $\tilde{A}=\{(v_1,v_2),(v_2,v_3)\}$。

当且仅当任意两结点由唯一的一条“路径”连接时，称一个图为一个“树”（tree）。图 5.4 中给出“树”和“非树”图的例子。

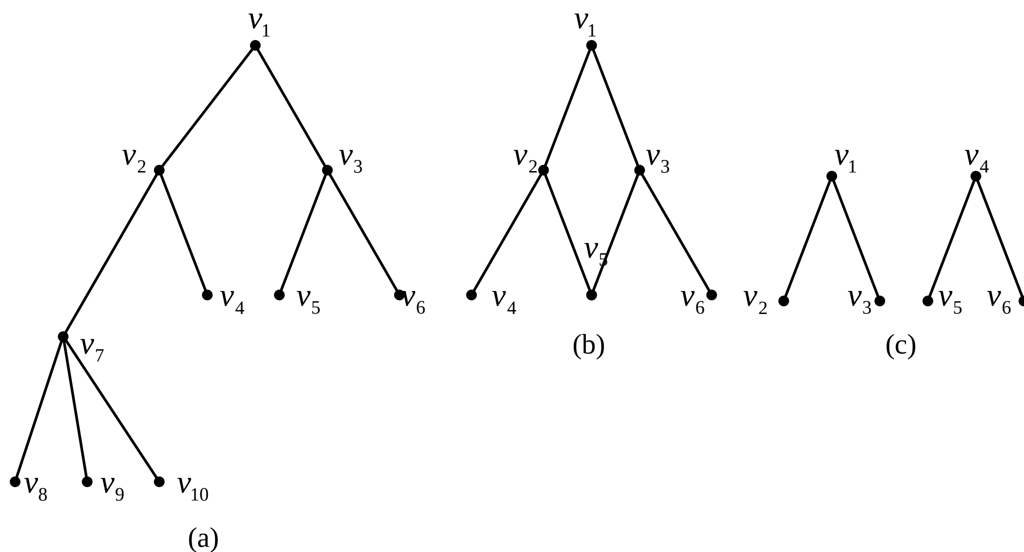


图 5.4　(a)中的图是“树”、(b)和(c)中的图不是“树”

显然，“树”中是没有圈的且任意两个结点间是“连通的”。

现假设图 T 是一个“树”，并设 r 为 T 中的一个结点，定义 r 为图的一个“根”（root）。我们现在可以将树中所有的“枝”进行“定向”，使得它们都是从“根”向外“定向”的。对于 $\widetilde{A}$ 中任意的 $(v_i,v_j)(i,j\in I)$，我们知道从 r 到 v_j 有唯一的一条路径。如果该路径通过结点 v_i，则枝 (v_i,v_j) 的方向为 $v_i\to v_j$ 或就记为 (v_i,v_j)；否则，从 r 到 v_i 的唯一路径[3]必经过 v_j，若不然，就不难找到两条从 r 到 v_i 的路径了，这违反了“树”的定义。因此，在这种情形可将枝 (v_i,v_j) 方向定向为 $v_j\to v_i$ 或记为 (v_j,v_i)。按照这种记法，我们知道 r 有如下性质：即 r 与另一不同结点 v_i 所构成的枝的方向必为 (r,v_i)，即从 r 只有“发出”（outgoing）的“枝”。相反，当从 r 只有发出的枝时，r 必为“根”。当结点 z 与树中任意其它结点 v_i 构成的枝的方向为 (v_i,z) 时，称 z 为“叶”（leaves）或终点结（terminal nodes），记 $Z=Z(T)$ 为由树 T 的所有终点结所构成的集合。图 5.5 给出一个树的例子。

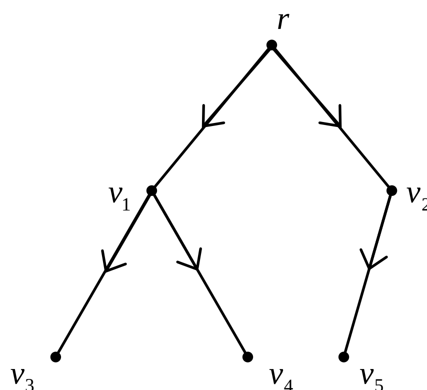


图 5.5　一个树的例子，其中 r 为根，$Z(T)=\{v_3,v_4,v_5\}$

注意，以上关于图及树的概念定义中不排除V（及A和$\tilde{A}$）中有无限个元素的情形。我们在今后在不引起混乱的情况下可省略去表示枝的方向的箭头。

有了上述概念准备，我们就可以正式给出 n 人动态博弈（扩展式表述的博弈）的定义。

定义 5.1 一个动态博弈（扩展式表述的博弈）G 被定义为同时满足下列条件的一个构成：

(i) $\Gamma=\{1,2,\cdots,n\}$是局中人集合，n 为自然数；

(ii) T 为一个树，称为博弈树，其中的根被称为初始结；

(iii) T 的非终点结集合 $T/Z(T)$被分成 n+1 个不相交的子集 $P^0,P^1,P^2,\cdots,P^n$。P^0 中的元素被称为“机会”（chance）或“自然”(nature)结；对于每个$i\in\Gamma$，P^i 中的元素被称为局中人 i 的决策结；

(iv) 对于 P^0 中的每个结点，从它向外发出的任一枝上被赋予了一个概率分布密度值；

(v) 对所有的 $i\in\Gamma$，P^i 又被分成 $K(i)$ 个互不相交的集合，称为“信息集”（information sets），记为$U_1^i,U_2^i,\cdots,U_{K(i)}^i$，满足：对于每个 $j=1,2,\cdots,K(i)$，有：

(a) 从U_j^i 中每个结点（$j=1,2,\cdots,K(i)$）发出的枝的数目相同，且U_j^i 中不同结点发出的枝的集合之间存在一个给定的 1-1 对应；

(b) 从根或初始结出发到终点结的每一个（定向的）路径至多通过每一个U_j^i 一次；

(vi) 对每个终点结 $z\in Z(T)$，对应于一个 n 维支付向量 $g(z)=(g_1(z),g_2(z),\cdots,g_n(z))$，其中 $g_i(z)$ 被称为局中人 i 的支付$(i=1,\cdots,n)$。

(vii) 上述从(i)到(vi)的完全描述是所有局中人之间的共同知识。

定义 5.1 在描述上较为抽象，但其好处是概念较为一般和严密。我们可以这样来直观理解其含义：每一个信息集中的结点是一个局中人进行决策的位置，称为决策结；博弈从树 T 的根开始。按归纳法表述，假定博弈已进行到非终点结v处。若v是局中人 i（即$v\in P^i$）的决策结，则局中人 i 根据v所在的信息集选择从v发出的一个枝。

由此，存在一条从初始结到某一终点结 Z 的唯一路径，博弈到 Z 处终止且每个局中人 i 获得支付 $g_i(z)$ $(i=1,\cdots,n)$；或者博弈从初始结开始沿着唯一一条路径进行下去但并无终点结，但仍然规定局中人 i 获得支付 g_i $(i=1,\cdots,n)$；前者称为“有限”次进行的动态博弈，后者称为“无限次”进行的动态博弈。

“自然”的选择通常表示局中人面临的不同的不确定性，是一种外生的条件。当不存在这种不确定性时，我们就得到完全信息博弈。

定义 5.2 在定义 5.1 中，当 P^0 为空集时，称 G 为完全信息动态博弈（Complete information dynamic game）。

定义 5.3 如果定义 5.1 中每个信息集$U_j^i(j=1,\cdots,K(i))$ 中只有一个元素（称为单结

信息集），则称 G 为完美信息博弈（Perfect information game）。完全信息博弈中没有外生的不确定性，而完美信息博弈则不存在任何的不确定性（事前和事后的不确定性皆不存在）。

5.1.2 完全信息动态博弈中的战略概念

局中人 i 的一个纯战略指对于局中人 i 在每一个信息集上所进行行动选择的一个规定。对局中人 i 来说，同一个信息集上的不同决策结应该在行动选择的范围上是不能有差别的，否则局中人就可以识别出他在同一信息集的哪一个决策结上，而这与信息集本来的含义相矛盾。对于局中人 i，我们用 $A_i(U_j^i)$ 表示其信息集与行动空间的对应，即 $A_i(U_j^i)$ 是局中人 i 在信息集 U_j^i 上所有决策结所能进行选择的行动的集合，称为局中人 i 在信息集 U_j^i 上的行动空间。

记 $I_i = \{U_j^i \mid j = 1, \cdots, K(i)\}$，即 I_i 为局中人 i 的信息集空间；定义映射 $S_i : S_i(U_j^i) \in A_i(U_j^i)$。则有

定义 5.3　称 S_i 为局中人 i 的一个纯战略。显然，局中人 i 的一个纯战略必是集合 $A_i = \prod_j A_i(U_j^i)$ 中的一个元素，反之亦然。

定义 5.4　称集合 $A_i = \prod_j A_i(U_j^i)$ 为局中人 i 的纯战略空间。

类似于战略式表述，我们可以有纯战略组合 $s = (s_1, \cdots, s_n)$ 及纯战略组合空间 $S = \prod_i A_i$ 等概念。同样，在完全信息动态博弈中我们也有类似于完全信息静态博弈中的纯战略纳什均衡定义，这里不再给出。

与静态博弈不同，在动态博弈情形，“战略”不等同于“行动”。在动态博弈中，在每一个信息集上有一个行动选择的问题，而一个局中人在由他进行行动选择的所有信息集所进行的行动选择构成他的一个战略，即战略是行动选择的一个谱系，一个战略规定了局中人在由他进行选择的所有信息集上所要选择的行动，即局中人在博弈开始之前所制定出的一个“相机行动计划”，它表明“如果……发生，我将选择……。”

在动态博弈中，一个战略规定了所有信息集上的行动选择，即使在博弈的均衡中某些信息集实际上达不到也如此。这是因为，也许正是那些在均衡中达不到的信息集上的特定行动选择规定才使得对应的均衡出现。譬如，在毛泽东的对外政治军事战略中，“人不犯我，我不犯人；人若犯我，我必犯人”就规定“我”在什么情形“不犯人”和什么情形“犯人”的行动选择。见图 5.6，其中当“人犯我”时，“我”（我国）选择“犯人”；当“人不犯我”时，“我”选择“不犯人”。

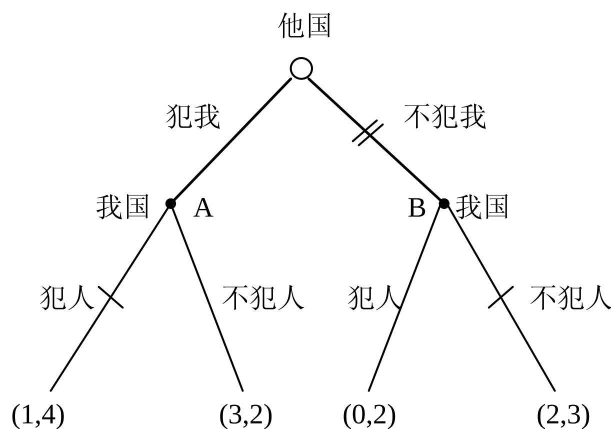


图 5.6　毛泽东的对外政治军事战略

显然，在信息集 *A* 处，“我”选择“犯人”，在信息集 *B* 处，“我”选择“不犯人”。我们用在枝上画上一条短线的方法来标识“我”的行动选择。给定“我”的上述战略，“人”或他国将选择“不犯我”的战略，我们用在枝上画两条短线的方法来标识“人”的行动选择。“人”之所以选择“不犯我”是因为他在初始结上进行行动选择时发现，如果他选择了“犯我”，则“我”必选择“犯人”，其所得支付为 1；而当他选择“不犯我”时，“我”必选择“不犯人”，其所得支付为 2；故他必选择“不犯我”。所以，“人”选“不犯我”，“我”选“不犯人”是一个均衡（在第三章已指出，可以用战略式表述给出扩展式表述博弈的另一种表达，而战略式表述博弈的纳什均衡已有定义，不难看出，这里给出的“均衡”正是一个纳什均衡），即可以作为博弈模型的一种预测结果。我们称枝“不犯我”和“不犯人”是均衡路径上的枝，它们组成的路径被称为“均衡路径”；而不是均衡路径的路径就被称为“非均衡路径”，如图中的由“犯我”和“犯人”两条枝构成路径（“犯我”和“不犯人”以及“不犯我”与“犯人”构成另外二条非均衡路径）。尽管信息集 A 不会达到，但“我”的战略规定“我”在 A 上选择“犯人”是起了作用的。这是因为，如果“我”在 A 上选择“不犯人”（作为理性人的“我”不可能作此选择），则“人”在初始结上就会选择“犯我”而不是选择“不犯我”了。

我们在第三章已经表明扩展式博弈可以用战略式表述方式给予表述，而在战略式表述博弈中，我们已经定义了纳什均衡、战略空间、支付函数、博弈结果等有关概念。这些概念可以直接转换成扩展式表述博弈中的对应概念，这里就不再重复给出这类定义了。

5.1.3 完全信息动态博弈的纳什均衡

我们在第三章中曾经指出，一个博弈的扩展式表述和战略式表述是等价的，只不过在动态博弈场合用扩展式表述更加直观一些。但有时通过战略式表述，我们可以在逻辑

上更加严格地分析动态博弈。也可以直接将战略式表述中已获得的结果直接应用于动态博弈，从而获得一些结果。譬如，根据第四章中的纳什存在性定理，我们直接得到如下结论：

定义 5.5　如果一个扩展式博弈有有限个信息集，每个信息集上局中人有有限个行动选择，则称该博弈为有限博弈[4]。

定理 5.1　如果一个扩展式博弈是有限博弈，则该博弈至少存在一个混合战略纳什均衡。

这里，扩展式博弈中的混合战略是从将扩展式博弈转换为战略式表述下的混合战略来理解的，即局中人从不同的相机行动选择计划中随机性地选择这些相机行动计划[5]。

我们用不同局中人在不同信息集上的行动选择来表示不同局中人的战略选择，并用不同局中人的战略选择组合来表示一个均衡，如图 5.6 给出的博弈均衡可表示为$\{(不犯我),(犯人,不犯人)\}$。其中大括号中第一个小括号表示第一个局中人的战略（第一个开始进行行动选择的局中人的战略），第二个小括号表示第二个局中人的战略；而在每一个小括号中，分别给出了局中人在不同信息集上的行动选择。其中，信息集的顺序是按照扩展式表述中从左到右编排的。当局中人和信息集的个数更多时，可依次类推。

在一些特别的场合，局中人面临由外部环境决定的不可控条件的变化（如天气变化和宏观经济条件变化），博弈论常将这种外生条件作为一种特别的局中人，称为“自然”（nature）。“自然”的行动选择对局中人来说事前都是未知的，但“自然”进行“行动”选择的信息集必定是单结的。对于所有信息集都是单结的扩展式博弈来说，我们曾将其定义为“完美信息博弈”。完全信息博弈与完美信息博弈是两个不同的概念。譬如，图 5.7 给出的博弈是完全信息博弈，因为其中没有“自然”这一个局中人，但却不是完美信息博弈。

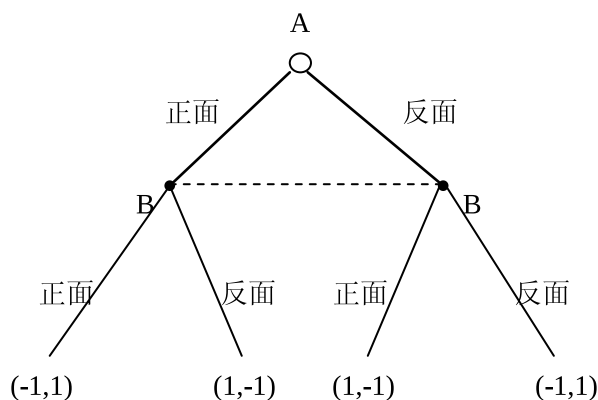


图 5.7　例 3.13 中“猜谜游戏”博弈的扩展式表述

在图 5.7 中，局中人 B 在决策时不知道局中人 A 会选择“正面”还是“反面”，故他不知道自己在决策时到底是位于自己信息集中的哪一个决策结上。B 的信息集是非单结的，所以不是完美信息博弈，但猜谜游戏却是一个完全信息博弈。在包括有“自然”这个局中人的博弈中，其他局中人通常不会知道“自然”的选择（至少不全知道，这来自于外生条件变化的不可预见性），至多只能知道“自然”以何种概率分布选择不同的“行动”（如知道天气变化的概率），这样，在包含有“自然”这个局中人的博弈中，信息就是不完全的。但我们在后面引入所谓“海萨尼转换”后，就可以将不完全信息博弈转化为完全但不完美信息博弈。这需要将“自然”这个局中人正式表达在博弈树中。

完美信息博弈对博弈中的有关条件作了相当严格的要求，因而是一种十分理想化的博弈。对于有限完美信息博弈，我们有如下定理：

定理 5.2 （Zermelo，1913；Kuhn，1953）一个有限完美信息博弈至少有一个纯战略纳什均衡。

为了完成定理 5.2 的证明，我们需要引入两个准备性的概念，即逆向归纳法（backward induction）和子博弈（Sub-game）。

先让我们来看看图 5.8 给出的博弈树。

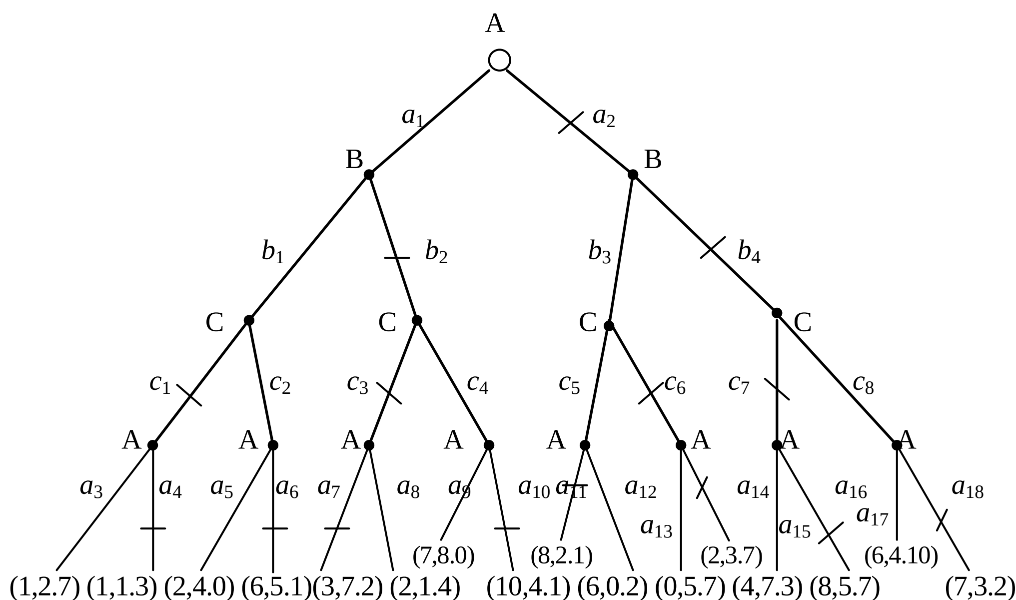


图 5.8　用逆向归纳法求解有限完美信息博弈的纯战略纳什均衡

这是一个有限完美信息博弈。所谓“逆向归纳法”，就是按与博弈顺序相反的方向逐一找出每一个局中人在其所有决策结上的最优行动选择，直到第一个博弈树的第一个决策结为止。由所有局中人最优行动选择构成的一条路径就是一个纳什均衡。首先，在倒数第一个进行行动选择的局中人顺序开始，我们在这个局中人 A 的所有倒数第一个决

策者上进行最优行动选择，并在所选出的最优行动a_4、a_6、a_7、a_{10}、a_{11}、a_{14}、a_{16}、a_{18}所对应的枝上画一个短线。

注意，a_4不是严格最优的行动选择。然后，我们再在倒数第二个行动顺序的局中人c的所有决策者上进行类似的最优行动选择，并在选出的最优行动c_1、c_3、c_6和c_7所对应的枝上画一条短线。按这种步骤，我们完成了剩下的 B 和 A 在第一次行动选择时的决策结上的最优行动选择，并在相应的代表最优行动的枝上画上一条短线。

根据已完成的各个决策结上的最优行动选择，我们来看看该博弈模型将会作出什么样的预测。首先，当博弈开始时，局中人 A 面临在两个行动a_1和a_2之间进行选择的问题。他知道，如果选择a_1则 B 将选择b_2，c将选择c_3，最后轮到 A 自己再度进行选择时，他将选择a_7，获得支付 3；相反，如果他开始选择的是a_2，则 B 将选b_4，c将选择c_7，最后轮到 A 自己再次进行选择时，他将选a_{16}，获得支付 8。因此，A 在博弈开始时会选择a_2，因为a_2是他在第一次进行行动选择时的最优行动。正如 A 所预料的是 B 随后将选b_4，然后是c选c_7，最后再由 A 选a_{16}，博弈结束。

这样，我们凭直觉认为行动组合（a_2，b_4，c_7，a_{16}）构成的路径是均衡路径，而纳什均衡就应该是战略组合$\{(a_2,a_4,a_6,a_7,a_{10},a_{11},a_{14},a_{16},a_{18}),(b_2,b_4),(c_1,c_3,c_6,c_7)\}$。其中，我们在 A 的战略表达中按其进行行动选择的顺序写出对应的行动选择，而在同一行动选择顺序中，我们按博弈树中从左到右的顺序写出对应的最优行动。

在完美信息博弈中，每一个信息集都是单结的。因此，如果我们设想博弈从任一个决策结开始，随后的过程仍然构成一个完美信息博弈。一般地，我们把从任一个单结信息集开始的作为原来博弈的一个组成部分称为一个“子博弈”，即当博弈进行到该单结信息集时开始的一个博弈[6]。

一般地，我们也可将原博弈当作一个特殊的子博弈。下面我们来证明完美信息有限博弈经逆向归纳法导出的战略组合必定是一个纯战略纳什均衡。显然，原博弈的一个战略组合限制在子博弈上给出子博弈上的一个战略组合。

定理 5.2 的证明：

我们采用归纳法和逆向归纳法来证明该定理。

当只有一个信息集时（由于这里考虑的是完美信息博弈，故此时只有一个决策结），局中人只需在其有限的行动空间中选择其支付最大的行动即可。由于是有限博弈，其行动空间是有限集，因而至少存在一个使其支付达到最大的行动。此时，这种选择也就构成了其支付最大的战略。当只有一个信息集时，该战略就是其纳什均衡战略。故此时逆向归纳法给出的战略组合是一个纳什均衡。现假定当信息集个数为 m 时定理 5.2 成立，往证信息集个数为 m+1 时定理 5.2 仍成立。

设信息集个数为 m+1 时博弈树的根或初始结为 r，则由初始结发出的枝为$(r,v_1),\cdots(r,\cdots v_h)$，从决策结$v_1,\cdots,v_h$出发又构成 h 个信息集个数不超过 m 的子博弈。显

然完美信息博弈的所有子博弈都是完美信息博弈，所以据归纳法假设，定理 5.2 对于这 h 个子博弈是成立的。现通过逆向归纳法对此博弈得到一个战略组合，它在这 h 个子博弈上给出在由逆向归纳法得到的纳什均衡，现在考虑第 i 个局中人在整个博弈中的最优战略。给定其他局中人的战略，当第 i 个局中人是初始结上的行动选择者时，他选择这样的行动即对应的发出枝为 (r,r_p)，使他在由 r_p 出发的子博弈上由原已存在的子博弈纳什均衡所获得的支付最大化。显然，这一战略就是他的最优战略。因为，给定其他人的战略，若他采用不同于此的战略，当他在子博弈上由原博弈战略所得到的子博弈战略不同于子博弈上原有的最优战略时，他的子博弈支付不会大于原有子博弈给出的支付，此时他在初始结上作任何选择都不会获得比前述战略更大的支付；当他在所有子博弈上由原博弈战略所得到的子博弈战略与子博弈上原有的最优战略相同时，他在初始结上作任何不同于前述战略的行动选择不会使他获得比前述战略更大的支付。显然，这一最优战略正是逆向归纳法给出的战略。

当第 i 个局中人不是初始结上的行动选择者时，由逆向归纳法给出的由 $v_1,\cdots,v_p$ 出发的所有子博弈上的战略组合显然据归纳法假设是纳什均衡，因为它在从 $v_1,\cdots,v_p$ 出发的每一个子博弈上都是最优的战略。因而第 i 个局中人的战略是给定其他人战略下的最优战略。

所以，由逆向归纳法得出的是纳什均衡，且定理证毕！

显然，定理 5.2 的证明过程实际上就是逆向归纳法构造式证明，它指出由逆向归纳法可得出有限完美信息博弈的一个纳什均衡。

我们实际上也证明，完美信息有限博弈经逆向归纳法给出的战略组合也给出了每一个局中人在每一个子博弈上给定其他局中人战略下的最优战略。

注意，同样是逆向归纳法也可能给出多个纳什均衡，见图 5.9。

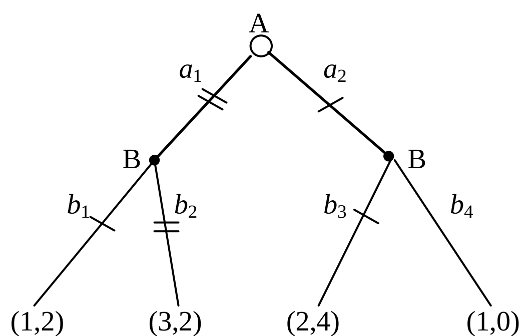


图 5.9　逆向归纳法给出两个纳什均衡

图 5.9 给出的博弈通过逆向归纳法获得两个纳什均衡，它们分别用单短线和双短线表示，即为 $\{a_2,(b_1,b_3)\}$ 和 $\{a_1,(b_2,b_3)\}$。

5.2 子博弈精炼纳什均衡

5.2.1 纳什均衡的缺陷与精炼

我们已看到，纳什均衡往往在同一个博弈模型中存在多重性，这对模型的可预测性造成问题。如果能够提出额外的理由将模型中出现的一些纳什均衡去掉，而留下更少的纳什均衡，这可能是博弈论专家所乐意做的事情。在博弈论研究中，一些学者总是在不断从逻辑甚至心理学的角度构造出这类理由，试图能将模型中出现的多个纳什均衡消去一些，从而提高模型的预测能力，这类工作被称为纳什均衡的“精炼”（refine）。在这里，我们介绍对动态博弈纳什均衡进行成功“精炼”的第一个工作，即子博弈精炼纳什均衡（Sub-game perfect Nash equilibrium）[7]概念，它是 Selten 完成的。

当我们分析静态博弈时，我们所说的战略是指局中人声明他将作何选择；如果博弈是动态的，则按照战略式表述，战略仍然具有这种含义。但是，当博弈是动态进行的时，与静态博弈相比，局中人就有了一种额外的选择，即事后机会主义。如果博弈是静态的，则局中人所声明的行动选择就是他们实际进行的选择，但是，如果博弈在行动选择上局中人有选择行动的先后顺序，那么，一些轮到稍后进行行动选择的局中人完全可以不按事前所声明的战略所规定的行动选择选择其行动，而是根据博弈进行到此时对局中人最为有利的方式选择行动。这就是说，在动态博弈中，即使局中人按事前所声明的战略组合构成一个纳什均衡，而这些均衡战略又规定了各个局中人在其所有信息集上的行动选择，这些行动选择也可能并非局中人在对应信息集上的最优行动选择。而当博弈实际进行到那些由纳什均衡战略规定的行动并非最优行动选择的信息集时，按照理性人假设，可以预言局中人届时不会按纳什均衡战略所规定的方式去选择行动，而是机会主义地选择最优的行动。这样，具有这种特点的纳什均衡就是不可信的，即不能作为模型的预测结果，按照“精炼”纳什均衡的思想，应当将其消掉。下面，我们给出几个例子来说明这种现象是可能出现的。

例 5.1 私奔博弈

姑娘爱上小伙，但其父反对，声称若女儿不与小伙继绝关系，他就要与女儿继绝父女关系。断绝恋爱关系使姑娘伤心，她的支付为 $(-a)$， $a>0$。而父亲很得意，其支付为 $b>0$；若已成婚，断绝父女关系使姑娘伤心，支付为 $-c$， $c>0$，而父亲默认已婚事实将使其难受，支付为 $-g$， $g>0$；但无论是否断绝父女关系，姑娘与小伙已成婚了，父亲的难受 $(-g)$ 总存在，但断绝父女关系失去女儿会加大父亲的伤心程度，其支付再加上一个 $(-d)$， $d>0$。如果父亲默认女儿嫁给的小伙子的事实，则女儿高兴，有正的支付 $e>0$。但父亲感到难受，支付为 $-g$。

但是，父亲的威胁是不可置信的。若已成婚，父亲的最优选择是默认事实，生米已做成熟饭$(-g > -d - g)$，所以，聪明的女儿一般不理会这类威胁（如汉初时卓文君与司马相如之私奔）。

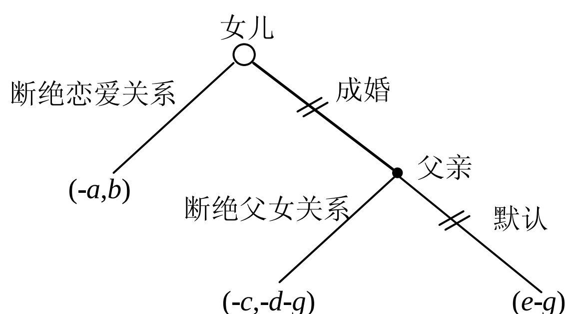


图 5.10　私奔博弈模型

例 5.2　有两个房地产开发商 A 和 B 在同时考虑是否在某地开发一栋新楼。若开发，则需 1 亿元投入。若他们各自都开发了一栋楼，每栋楼售价为 0.7 亿元；若只有一个开发商开发，每栋楼售价为 1.1 亿元。

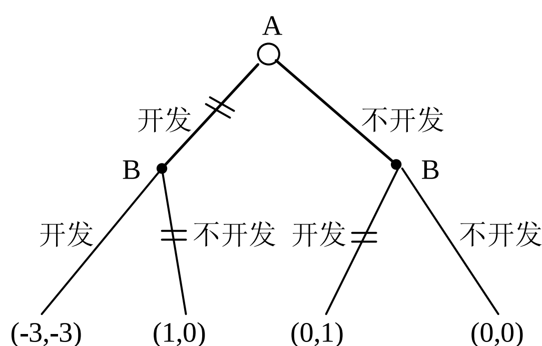


图 5.11　房地产开发博弈

在图 5.11 中，我们用千万元作为支付单位。

这个博弈有三个纳什均衡：$\{$不开发,(开发,开发)$\}$，$\{$开发,(不开发,开发)$\}$，$\{$开发,(不开发,不开发)$\}$。为了看出这一点，只要按表 3.2 那样的方式将博弈写成战略式表述即可，见表 5.1。

表 5.1　房地产开发博弈的战略式表述

		B			
		(开发,开发)	(开发,不开发)	(不开发,开发)	(不开发,不开发)
A	开发	-3,-3	-3,-3	1,0	1,0
	不开发	0,1	0,0	0,1	0,0

通过划线法，我们从表 5.1 中得到上述三个纳什均衡。显然，{不开发,(开发,开发)}含有不可置信的威胁，因 B 威胁无论 A 是否开发，他都将开发，但当 A 果真选了“开发”，B 却会选“不开发”而不是“开发”。均衡{开发,(不开发,不开发)}含有不可置信的承诺，即当 A 不开发时 B 也不开发，但当 A 果真不开发时，B 却会选“开发”而不是“不开发”。所以，只有{开发,(不开发,开发)}才是合理的均衡，可作为模型预测结果，它就是子博弈精炼均衡。

5.2.2 子博弈精炼均衡概念与求解

从例子中看出，在动态博弈中，因为一个战略是对局中人所有自己行动的信息集上行动选择的一个规定（计划），而纳什均衡是局中人之间的这种规定（计划）之间的均衡，但其中可能含有“不可置信的规定”（如不可置信的承诺或威胁），这种不可置信的规定会导致局中人的事后机会主义行为，即届时轮到自己行动了，局中人会根据当时的情况寻求最优行动选择，而这种最优行动选择可能违反事先战略规定中所“规定”的在此时所应作出的选择。

理性的其他局中人会事先明白这一点，会预知到这种机会主义行为，因而不会相信含有“不可置信规定”的战略。因此，由那些含有“不可置信规定”战略所构成的纳什均衡是不能作为模型预测结果的，必须去除掉。

这就引出了“子博弈精炼纳什均衡”的概念，即只保留那些在每一个信息集上都是最优（纳什均衡）行动的战略所构成的纳什均衡，对于有限完美信息博弈来说，它通过逆向归纳法就可求出。

对于有限完美信息动态博弈，如果通过逆向归纳法求得其纳什均衡，则这种纳什均衡显然去掉了所有不可置信的规定，从而是子博弈精炼纳什均衡。Selten(1965)引入的“子博弈精炼纳什均衡”使纳什均衡第一次获得精炼，使模型中的纳什均衡减少一些，但仍不能保证均衡是唯一的。

子博弈精炼纳什均衡的概念思路是要求局中人的行动选择在所有信息集上都是最优的，这就消除了不可置信的战略。该概念的定义为：

定义 5.6 扩展式博弈的战略组合 $s^* = (s_1^*, \cdots, s_n^*)$ 是一个子博弈精炼纳什均衡（以后可简称子博弈精炼均衡），如果：(1)它是原博弈的纳什均衡；(2)它在每一个子博弈上的限制都给出了子博弈上的一个纳什均衡。

由于子博弈精炼纳什均衡要求战略组合在所有子博弈上都构成纳什均衡，故不存在任何偏离事前战略的事后机会主义行为，实际发生的动态行为与事前战略规定是一致的，博弈模型的预测是可信的。为什么说在子博弈上是纳什均衡呢？因为合理的均衡要求战略在所有信息集上都是最优行动选择，当所有局中人的行动选择都是最优的时，就构成纳什均衡[8]。

一般地，对于有限的完美信息动态博弈，我们可以运用逆向归纳法求出子博弈精炼纳什均衡，该方法实际上是重复剔除劣战略方法在扩展式博弈中的应用，是从倒数第一个决策结开始剔除“劣行动”，然后依次倒推至初始结。但是，对于无限博弈和不完美信息博弈，逆向归纳法不适用。因无限博弈或者无倒数第一个决策结，或者某决策结上有无限个行动（难以选优）；而不完美信息博弈中存在包括多个决策结的信息集，由于局中人不能区分它们，故不能定义“最优行动”。尽管如此，有时也可用逆向归纳法的思路去找出不完美信息博弈的均衡解。

另外，无限博弈有时也有均衡解，只不过需用其它方法求解。

5.2.3 子博弈精炼纳什均衡的缺陷

正如我们在第三章中已指出的那样，由于重复剔除劣战略的占优战略均衡对局中人的理性要求太高，以致于均衡解有时令人感到难以实际出现，或者说重复剔除劣战略的占优战略均衡对局中人的理性程度要求太高，使得人们对模型可预测的结果是否真正就是均衡所预言的那样感到怀疑。因为通过逆向归纳法所得到的完全信息动态博弈子博弈精炼纳什均衡实质上是在每一个信息集上进行了重复剔除劣战略的过程。因而子博弈精炼均衡实际上也存在如重复剔除劣战略的占优战略均衡那样的一些缺陷或局限，如对局中人的理性程度要求过高。

例 5.3

图 5.12 中的完全信息动态博弈有一个子博弈精炼均衡：

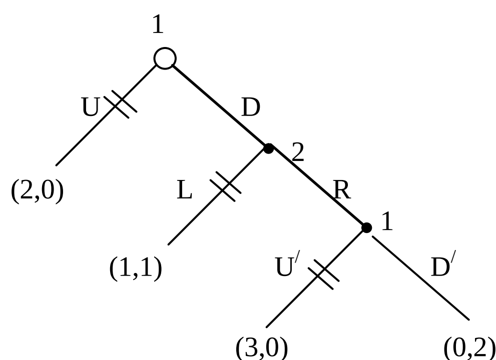


图 5.12　对局中人理性程度要求过高的子博弈精炼均衡

即使两个局中人都是理性的，但如果 1 认为 2 是非理性的（认为 2 是傻子），则 1 在第一阶段可能选 D，期待 2 在第二阶段选 R，从而自己在第三阶段选 U' 得到 3 单位支付。或者，即使 2 知 1 是理性的，但如果 1 不认为 2 会相信自己是理性的，1 就可能在第一阶段选 D，期待 2 认为自己不是理性的（认为 1 是傻子）因而在第二阶段选 R 期待自己在第三阶段选 D'[9]。

为了说明子博弈精炼均衡的这一局限，我们还可以在更多阶段的动态博弈中更明显地看出这一点。在例 5.4 中，我们有一个 n 人完全信息动态博弈，而博弈也由 n 个阶段构成。

例 5.4

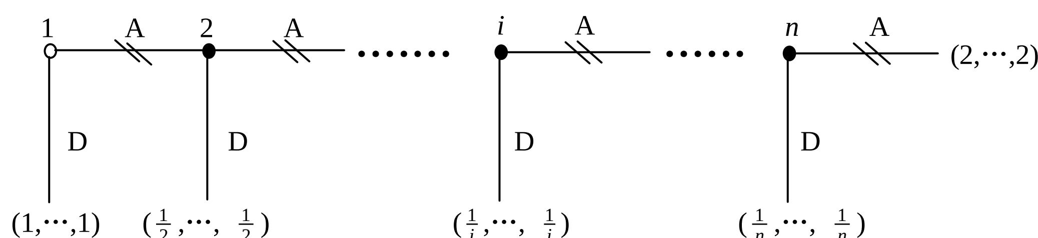


图 5.13　n 人 n 阶段完全信息动态博弈

子博弈精炼均衡为 $(A,\cdots,A)$，即所有局中人都选 A。但当 n 很大时，这一预测就值得怀疑了。对 1 来说，为了获得支付 2，要求所有其他的 n-1 个局中人都要选 A。当 1 不敢确信其他 n-1 个局中人都一定会选 A 时，他就可能选 D 而不是 A。譬如，若一个局中人选 A 的概率为 P<1（局中人在选择时可能犯错误），则 n-1 个局中人都选 A 的概率为 P^{n-1}（假定不同局中人犯错误的概率是相互独立的）。即使 P 很大（只有 P<1），P^{n-1} 也会很小（当 n 很大）。

另外，即使 1 确信其他 n-1 个局中人都会选 A，他也不能保证 2 是否也相信后面的 n-2 个局中人都会选 A。当 1 认为 2 不相信后面 n-2 个局中人都会选 A 时，1 会认为 2 会选 D。基于这种推理，从而 1 在开始就会选 D。

逆向归纳法要求支付向量是所有局中人的共同知识，即“1 知 2 知 3 知……的支付向量”。这一要求太强，超乎现实。当 n=3 时，要求 1 和 2 知 3 的支付，而且还要求 1 知 2 知 3 的支付，只有如此，1 才能预测 2 对 3 的选择的预测。若 1 认为 2 不能正确预测 3 的选择时，1 就会选 D。

5.2.4 对子博弈精炼纳什均衡预测偏差的解释

由于子博弈精炼均衡具有上述对局中人理性要求过高的特点，因而运用这种均衡概念对模型结果进行预测可能就会出现偏差。当实际观察到的行为与子博弈精炼均衡预测

的行为之间出现偏差时，博弈论专家对此有几种不同的解释或处理办法。譬如，我们在下面以例 5.5 为例子来说明。

例 5.5　Rosenthsal(1981)模型

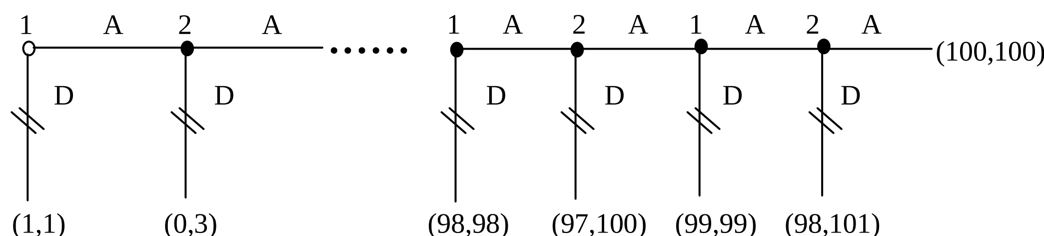


图 5.14　Rosenthsal 博弈模型

唯一的精炼均衡为（*D*,……,*D*），1 在初始结上选 *D* 而结束博弈。每人得支付 1。但若都选 *A*，则各得支付 100。这里，较高的理性要求反而是不好的。当 2 知 1 是理性的，2 也知 1 知 2 是理性的，2 会确信 1 会在开始就选 *D* 而结束博弈。但若降低对“理性”的要求，比如在实际中 2 发现 1 实际上选了 *A* 而不是预测中的 *D*，此时 2 又该如何进行下面由他作出的选择呢？根据逆向归纳法，2 应选 *D*，因若 1 得到下一个机会的话，1 会选 *D*，这里实际上假设 2 认为 1 在开始选 *A* 是“偶然犯了一次错误”，而 2 选 *D* 正是在此假设下的正确反应。

一般地，当 2 没有预料到的情况发生时，2 的最优选择依赖于 2 如何预测 1 未来的行为，特别是 2 如何修正其对 1 理性程度的评价或者 2 认为 1 对 2 的理性程度的评价，即 1 开始就选 *A* 可能是因他是非理性的（1 是傻子），也可能是因 1 认为 2 是非理性的（1 认为 2 是傻子）。

若 1 不是理性的，或若 1 认为 2 是非理性的，或者若 1 认为 2 认为 1 是非理性的，2 就会选 *A*。反之，1 会如何考虑 1 选 *A* 对 2 的影响呢？1 可能会故意选 *A* 以诱使 2 认为 1 是傻子因而也选 *A*，期待这个“将错就错”过程一直进行下去以获得 100 单位支付（见注释[9]）。

现在的问题是：2 如何从 1 选了 A 这一事实就推出 1 是非理性的呢？逆向归纳法没有为当未预料到的事情出现时局中人如何形成他们的预期提供解释，使得逆向归纳法的逻辑受到怀疑。下面是不同的专家对此作出的解释。

①FKL 理论

Fudenberg，Kreps 和 Levine（1988）提出：当博弈出现偏离原博弈子博弈精炼均衡时，说明原博弈设定的“支付函数”不同于实际上真实的支付函数，即有关“支付函数”的信息不确定造成这种偏离行为。因此，局中人一旦观察到未曾预料到的行为，他就应

该修正有关支付函数的信息。他们证明：任何观察到的行为都可用博弈对手的某种特定支付函数来解释，从而将偏离行为出现后如何预测博弈结果的问题归纳为选择哪一个支付函数的问题，从而回避了当零概率事件出现时如何形成新的信念的困难，也即修正支付函数后就接此新的支付函数计算新的子博弈精炼均衡就是观察到偏离行为后的局中人要做的事情。

②颤抖手理论

Selten（1975）为捍卫他本人提出的子博弈精炼均衡和逆向归纳法，将偏离行为解释为局中人在博弈中犯的错误，即均衡的“颤抖”（trambles）。他指出：扩展式博弈允许局中人在实际选择行动中犯错误。如果局中人在每个信息集上犯错误的概率是独立的（因而局中人不会犯系统性的错误，今后不会常再犯错误），则无论过去的行为怎样，局中人在今后应继续使用逆向归纳法预测从现在开始的子博弈的行为（既往不咎）。

5.3 博弈中的承诺行动

5.3.1 什么是承诺行动

一个博弈模型的均衡显然与博弈模型中各个局中人的战略空间或行动空间有关，当我们改变博弈模型中一些局中人的战略空间或行动空间时，也就改变了博弈的均衡。如果一个局中人对已有博弈的均衡感到不满意，他（她）可以通过改变其行动空间从而导致其战略空间发生相应的变化来避开他不满意的均衡并获取其所需的均衡。当在一个博弈中，倘若某局中人希望一个本属含有不可置信威胁或承诺的行动的非精炼均衡能真正发生，他可以通过改变其行动空间（通常是减小其行动空间或战略空间）使其威胁或承诺变得不可置信，从而将原本非精炼的均衡变成精炼的均衡（战略或行动空间改变后，博弈本身也改变了），这类局中人改变其行动或战略空间的行为被称为“承诺行动”（Commitment action）。

在例 5.2 中，精炼均衡是 *A* 开发 *B* 不开发，另一个纳什均衡即 *B* 不管 *A* 是否开发都要开发和 *A* 不开发被作为非精炼均衡剔除掉了。这是因为，该均衡中含有不可置信的威胁，因为一旦先行动的 *A* 开发出一栋大楼，理性的 B 就不会如他事先威胁的那样会开发。但是，倘若 *B* 希望这个非精炼均衡能真正出现，即他自己开发而对方不开发（显然，*B* 会希望如此的），他就可以通过与一个第三者签一份合同来达到这一目的。该合同要求 *B* 无论 *A* 是否开发也应开发一栋大楼，否则 *B* 向第三者 *C* 支付 4 千万元（*B* 与 *C* 打一个赌）。这样一来，*B* 必定在任何情形下（无论 *A* 是否开发）都会开发的，因为即使 *A* 开发，*B* 开发也只损失 3 千万元，不开发会损失更多的 4 千万元。注意，这样的合同一定要让 *A* 看见知悉。这样一来，*A* 就会选择不开发了，而 *B* 单独开发出一栋大楼并赢利 1 千万元。在这个新的均衡中，*B* 让原本是非精炼的均衡（不开发，(开发，开发)）变成

了子博弈精炼均衡，B 不仅未支付 C4 千万元而且还赢利 1 千万元。图 5.15 给出了这一新博弈及其子博弈精炼均衡，其中的承诺行动就是 B 与 C 进行的打赌或所签的合同。

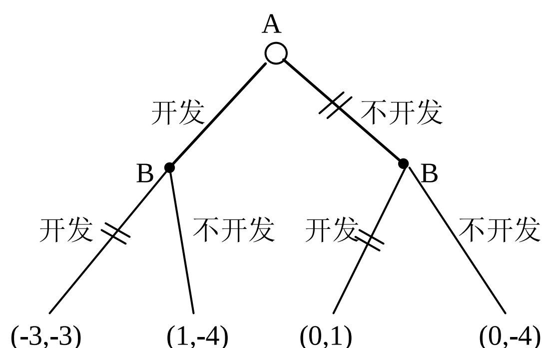


图 5.15　承诺行动下的房地产开发博弈

承诺行动是局中人通过减少自己在博弈中的选择机会来迫使对手选择自己所希望的行动。当局中人选择使自己选择某些行动的成本增至无穷大时，称为“完全承诺”；当只增大自己选择某些行动的成本时，称为“不完全承诺”。上述例子就是一个不完全承诺，开发商 B 通过增大自己选择不开发的成本来迫使 A 选择“不开发”，而项羽的“破釜沉舟“则是一个完全承诺的例子。当然，完全与不完全承诺的划分是相对的，选择特定行动的成本变得很大时，不完全承诺实际上就成为了完全承诺。

5.3.2 真实的故事：承诺行动的例子

例 5.6 承诺行动——项羽的破釜成舟

秦朝末年，反秦义军在新上任的统帅项羽的率领下，渡过大河与秦军精锐主力决战。当时的情况是，秦军主力是由名将章钳率领的精锐之师，而项羽统领的义军是一群缺乏训练，给养不足的乌合之众，且项羽本人又是刚刚通过斩了统帅宋义而自任统帅上台的，军心欠稳。两军相比，秦军无论在人数、装备及给养，还是士兵素质方面都远强于义军。若此两军相遇，好似狮犬之搏，一般人都会认为义军不是章钳大军的对手（宋义是在义军统帅项梁去世后接替项梁而出任义军统帅的，但在决战前夕就是因此顾虑而终日饮酒不敢出战被愤怒中的猛士项羽所杀）。这样，决定战争胜负的因素就取决于两军的士气了。项羽这个粗人是深知这一点的，他在义军渡过河后令人击碎煮饭的大锅（破釜），还将渡河用的船只悉数尽沉河底（沉舟），然后告诉义军士兵：“我们已没有退路了，只有不顾一切地猛击秦军，才有一线生路”。结果义军果真一鼓作气大败秦军，俘虏了秦军大将章钳。此战实际上为彻底推翻秦王朝打下了基础，从此义军一路顺风地打到了秦朝国都咸阳[10]。

用博弈论的语言来描述历史上的这一著名战役，我们说项羽的破釜沉舟就是一个

“承诺行动”。对于义军士兵来说，其行动空间在项羽破釜沉舟之前可以说有四个元素，即｛勇猛进攻，与秦军僵持不下，投降秦军，乘船返回逃跑｝。如果两军相遇，义军选择“勇猛进攻”会冒很大风险，因为秦军太强大了；如果义军选择僵持不下，也不是个办法，因为拖延进攻时间对义军并无好处；如果义军选择投降，则按当时的情况无异于自取灭亡，因为秦军肯定会对投降的义军赶尽杀绝（当时的战争并无“优待俘虏”的说法）。因此，乘船逃跑很可能就是义军在强敌面前会出现的结果了。同时，秦军知道义军有退路，因而预料义军可能是一触即散的乌合之众，不会遇到顽强抵抗，因此他们会勇猛作战。再回过来，义军知道秦军有如此心理，更加对秦军产生了畏惧，选择逃跑可能是更应考虑的退路了；给定义军的这种心理状态，秦军会进一步增强信心，因而进攻会更加有条不紊和猛烈；而给定秦军的这种心理，义军就会更加胆怯，……，如此往复，可以猜想最后的结果怎样——几乎可以肯定义军会在战斗开始就出现混乱，而秦军如虎狼之师般横扫义军于河滩上。

然而，当项羽作了“破釜沉舟”的承诺行动之后，义军士兵的行动空间就减少了三个元素，变为｛勇猛进攻，僵持不下，投降｝。两者比较，“勇猛进攻”是占优于“投降”和“僵持不下”的，因为根据当时的情况，投降无异于自取灭亡，而选择“勇猛进攻”还可能打败秦军而获取一条生路。同样，僵持不下只会增加对方的实力，因为己方无退路而对方有后援。因而义军在此情形下必定会选择万众一心的勇猛杀敌战略。给定义军的这种选择，秦军反而胆怯了，因为他们遇到了义无反顾的拼死大军；而给定秦军胆怯，义军在心里又增强了战胜敌军的勇气；而给定义军的这种勇气，秦军会进一步准备溜走逃命而不打算拼死作战，……，如此反复，我们看到“破釜沉舟”这一承诺行动彻底扭转了两军在心理上从而在士气上的对比情况，因而使义军在人数、训练和装备较敌军为劣的情况下，通过在心理士气上占优势而取胜。

破釜沉舟是战争史上运用承诺行动的一个著名战例，在其他的古代战争故事中也不乏此类例子，如韩信赵国之战时，将被赵军追击下的大军故意引至绝无逃路的大海边，然后高呼我们无退路了，只有拼死一战才有逃生之望，结果绝望中的士兵拼死反击追军而大获全胜。事后韩信称此计为“置之死地而后生”。其它还有三国时的曹操与袁绍的仓亭之战等。有鉴于此，古兵书中有明训“穷寇勿追”[11]，以免被追急了的对方反咬一口。

例 5.7 企业的过剩生产能力

经济学家发现，在许多行业中，都存在过剩生产能力的现象；同时，特别是在新兴行业中，一些先进入的企业在并不知晓未来市场大小的情况下，一味地建造大规模的生产基地和安装生产装备。这是为什么呢？博弈论专家对此给出的解释是，企业为了阻唬潜在的竞争对手，通过显示其过剩生产能力来给潜在竞争对手一个“可置信的”威胁：你要是进入行业与我竞争，我并不会减少产量。这样，企业保持过剩生产能力就是一种

“承诺行动”，其原理如下：

我们知道，任何行业中的企业生产都存在一个所谓的“盈亏平衡点”，即当产量小于某一阈值时，企业就会出现亏损；当产量大于该产量阈值时，企业才会有利润可获。这个产量阈值就被称为企业的“盈亏平衡点”。不同行业的盈亏平衡点是不相同的，一般地，生产规模或生产能力愈大（资本密集型）的企业，其盈亏平衡点就愈大，因为对于资本投入较多的企业来说，只有较大的产量才能将较大的固定资本分摊到较多的产品上去，从而平均成本较低，利润就较高。

当一个行业中有一家企业拥有较大的生产能力时，如果其它的企业进入该行业与之竞争，该企业不会大幅降低产量。这是因为，给定有限的市场容量，新的企业进入后，如果原有企业不降低产量，向下倾斜的需求曲线就会将产品价格拉低，因为市场上的产量增加了。如果原来的企业生产能力较小，其盈亏平衡点也较小，它减少产量不一定导致亏损，因而原有企业在新的竞争性对手进入后会减少产量，尽管它可能事先威胁它不会减产，因为这种威胁是不可置信的。但是，当原有企业生产能力较大时，其盈亏平衡点就较大，其威胁就是可置信的了。因为对于有较大盈亏平衡点的企业来说，维持一定的较大产量是企业不亏损的基本条件。此时，新的企业进入后，原有企业不会有大幅减产，而新企业也有一个基本的盈亏平衡点，因而市场上就会有较多的产品，产品价格就会有明显下降，结果可能使进入的企业赚钱不多甚至亏损。由于进入行业有一次性的进入成本，所以在这种情况下进入企业就很可能导致亏损。如果事先进入企业就明白这一点，它就不会进入了，而原有企业正是通过其维持的高生产能力，从而成功地将潜在的进入者拒之行业外，维持长期的垄断性经营。

对于新兴行业，重要的是先进入并成功地将后来者阻隔于行业之外，因而许多企业在进入新兴行业时首先是扩大规模（尽管它们并不知道新兴市场的容量有多大），通过这种承诺行动将后来的竞争者拒之门外。

企业的过剩生产能力往往就是垄断性企业为了达成高生产能力而过度投资的结果，因为它们很难事先准确预测市场容量；同时，过剩的生产能力正好也向潜在竞争者显示了自己是高生产能力的企业（否则谁会相信你曾作了承诺行动的呢？），这样可使对竞争者的威胁变得可置信[12]。

5.4 几个应用例子

本节将介绍几个子博弈精炼纳什均衡的应用性例子。这些例子是应用子博弈精炼均衡的分析方法研究重要的经济理论疑难问题中的成功案例，由此我们也可以领略到运用博弈论方法进行经济分析的一些绝妙之处。

为了以下有关 Stackelberg 博弈的应用例子，这里先介绍一下什么是 Stackelberg 博弈。

5.4.1 Stackelberg 博弈

我们将逆向归纳法的思想扩充于企业在市场上进行竞争的产量战略上去，就得到所谓 Stackelberg 博弈。在企业进行产量决策有先后顺序的情况下，博弈自然是动态的，但由于企业的产量通常是一个连续变量，因而博弈本身不是有限博弈，但仍可通过逆向归纳法的思路导出其子博弈精炼均衡。

假定有两个企业在同一产品市场上进行竞争，但它们的产量决策在时间上有先后之分，且后行动者能观察到先行动者的产量决策结果。设第 i 个企业的产量为 q_i，利润为 $\pi_i = \pi_i(q_1, q_2)$，其中 $q_i \in [a_i, b_i],\ i = 1,2$。$i = 1$ 是先行动者，$i = 2$ 在看到 1 的产量 q_1 后决定其产量 q_2。1 知道它每决定一个产 q_1，2 的产量决策将为 q_2，满足：

$$q_2 = q_2(q_1) \in \arg\max \pi_2(q_1, q_2) \qquad (5.1)$$

即给定 q_1，q_2 必是 2 的子博弈占优战略。

所以，1 的选择就是 q_1，它满足：

$$q_1 = \arg\max \pi_1(q_1, q_2(q_1)) \qquad (5.2)$$

即给定式（5.1）确定的反应函数 $q_2(q_1)$，1 选择其占优战略 q_1^*，它由式（5.2）决定，2 选 $q_2^* = q_2(q_1^*)$。

显然，这样得到的解 (q_1^*, q_2^*) 必定是子博弈精炼均衡。称这种博弈为 Stackelberg 博弈（Stackelberg game）。Stackelberg 在 1934 年提出这一概念（Stackelberg，1934），其均衡概念在子博弈精炼纳什均衡（1965）出现以前就有了，但它可被视为精炼均衡的最早版本。与古诺模型不同，这类模型中的企业 1 是领导者（leader）、企业 2 是追随者（follower），前者先选择产量 q_1，后者在观察到 q_1 后选择自己的产量 q_2，两者不是同时行动的。这是一个完美信息动态博弈，但战略空间是连续性空间。下面，我们利用例 3.10 中的企业利润函数、成本函数及需求函数来计算一个 Stackelberg 博弈的子博弈精炼均衡。此时逆需求函数为 $p = a - q_1 - q_2$，两个企业有相同的不变边际成本 $c \geq 0$，企业 i 的支付（利润）函数为：

$$\pi_i(q_1, q_2) = q_i(p - c),\ \ i = 1,2$$

据式（5.1）有：$q_2 = q_2(q_1) = \dfrac{1}{2}(a - q_1 - c) \quad (\text{设} q_1 < a - c)$

$$q_1^* = \frac{1}{2}(a - c)$$

代入 $q_2(q_1)$ 得 $q_2^* = \dfrac{1}{4}(a - c)$

子博弈精炼均衡结果为 $\left\{\dfrac{1}{2}(a - c), \dfrac{1}{4}(a - c)\right\}$。

将此例与前述例 3.10 及例 3.26 中的古诺模型相比较。古诺博弈均衡结果为

$q_1^* = q_2^* = \frac{a-c}{3}$，而 Stackelberg 博弈均衡结果为 $q_1^* = \frac{a-c}{2} > \frac{a-c}{3}$，$q_2^* = \frac{a-c}{4} < \frac{a-c}{3}$，即领导者的产量变大了，追随者的产量变小了（不难证明，领导者的支付变大了而追随者的利润却减小了）。这是因为博弈中存在"先动优势"（不难证明当企业进行价格而非产量竞争时，反而是"后动优势"）。

同时，该例还说明，拥有信息优势可能使局中人处于劣势（后动者拥有更多信息），而这在决策论中是不可能的，这是博弈论与决策论的不同。企业 2 处于劣势是因为它在行动前已知企业 1 的产量，而企业 1 在开始行动时也知如此。但若企业 2 不知企业 1 的产量且企业 1 也知如此，则即使企业 1 先行动，博弈也是古诺的而非 Stackelberg 的，企业 2 反而获益，企业 1 的先动优势就不存在了。

企业 1 先生产产量就是一种承诺行动，生产出来的产量是沉淀成本，从而使企业 2 不得不认为它的威胁是可置信的。如果企业 1 只是宣布它将生产 $q_1^* = \frac{a-c}{2}$。企业 2 不会相信它的威胁，因若企业 2 相信它的威胁而选 $q_2 = \frac{a-c}{4}$，给定此 q_2，企业 1 的最优选择是 $q_1 = \frac{3(a-c)}{8}$ 而不是 $q_1 = \frac{a-c}{2}$（不难证明）。见图 5.16。

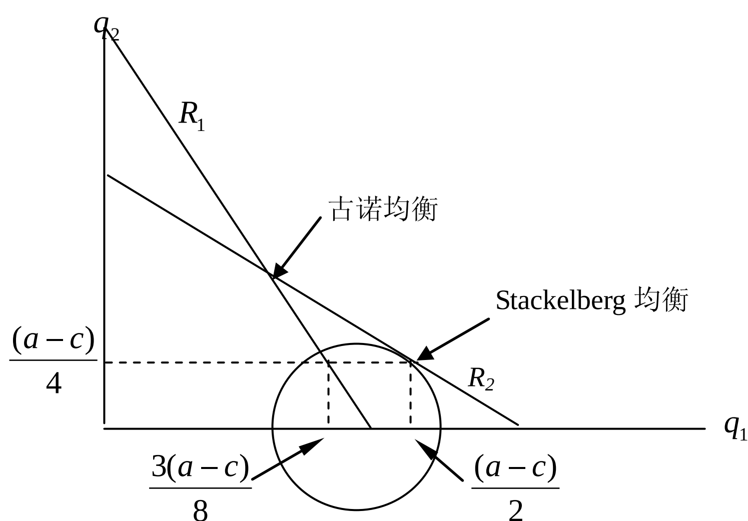


图 5.16　Stackelberg 均衡与古诺均衡的比较

例 5.8　里昂惕夫工会模型

经济学家 Leontief 于 1946 年提出的工会模型描写了工会与厂商就工资进行谈判的机理，并就均衡的非帕累托最优性和谈判结果的不稳定性作出解释，它解释了人们所观察到的工会与厂商就工资水平所进行的经常性谈判是如何发生的（Leonfief, 1946）。

局中人 1——工会：战略空间为$\{提出工资水平w\}$

局中人 2——厂商：战略空间为$\{决定工人雇用量L\}$

工会的支付函数为$U(w,L)$，假设$\dfrac{\partial U}{\partial w}>0$，$\dfrac{\partial U}{\partial L}>0$，厂商利润函数或支付函数为$\pi(w,L)=R(L)-wL$，其中：$w$为工资水平，$L$为工人雇用量，$R(L)$为厂商收益函数（销售值），假设$R(L)$是$L$的增函数且为凹函数。

博弈的行动顺序：

(1)工会提出工资水平；

(2)厂商观测到（并接受）w，随后选择雇用人数L；

(3)工会和厂商的支付分别为$U(w,L)$和$\pi(w,L)$。

下面，见图 5.17。

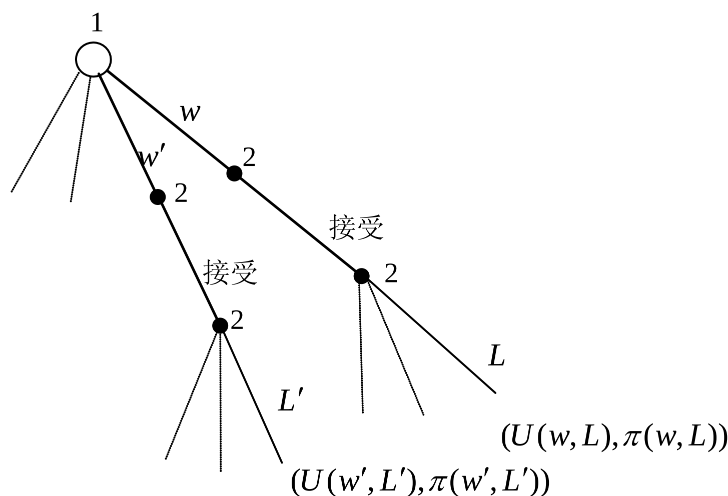


图 5.17　里昂惕夫工会模型的博弈树

下面，我们运用逆向归纳法求解此博弈。

对于第一阶段中工会提出的任意的工资率w，厂商在第二阶段中通过最大化其利润来决定最优的劳动量需求$L^*(w)$，即

$$\max_{L\ge 0}\pi(w,L)=\max_{L\ge 0}\{R(L)-wL\}$$

一阶条件为$R'(L)=w$，解出$L^*=L^*(w)$

为保证一阶条件在$w>0$时总有解，假定$R'(0)=\infty$，$R'(\infty)\ge 0$，即$R'(L)\in[0,\infty)$。

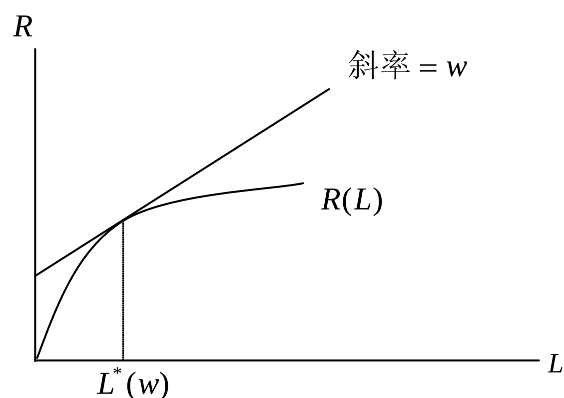


图 5.18　收益函数的斜率等于工资率

给定 w，利润最大化给出厂商最优雇用工人量 $L^{*}(w)$，当 w 上升时，工人雇用量减少（见图 5.18），故 $L^{*}(w)$ 是 w 的减函数（见图 5.19）。

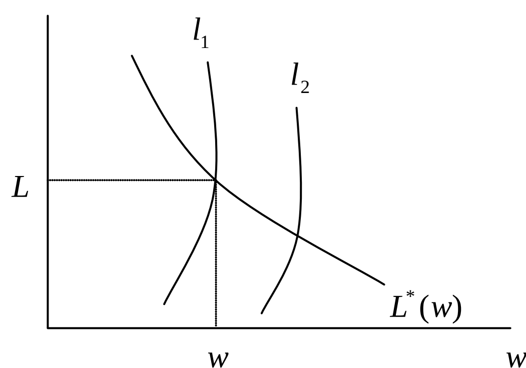


图 5.19　凹向纵轴的等利润线

给定 w，由 $L^{*}(w)$ 给出的 L 是利润最大化工人雇用量，在图 5.19 中，等利润线必凹向纵轴。因为等利润化满足 $\pi = R(L) - wL$，给定利润水平 π，等利润线方程为 $w = \frac{R(L)}{L} - \frac{\pi}{L}$，故有

$$\frac{dw}{dL} = \frac{d\left[\frac{R(L)}{L}\right]}{dL} + \frac{\pi}{L^2}$$

当 $L \to \infty$，因 $\frac{d\left[\frac{R(L)}{L}\right]}{dL} < 0$，故 $\frac{dw}{dL} < 0$

当 $L \to 0$，$\frac{\pi}{L^2} \to +\infty$，故 $\frac{dw}{dL} > 0$

所以，等利润线凹向纵轴。显然，愈是位于右边的等利润线，利润就愈低，因它对应同样的 L 却付出更高的工资，或同样的 w 却雇用不同于 $L^{*}(w)$ 的工人量，由于 $L^{*}(w)$ 是利润最大化工人雇用量，故 $L^{*}(w)$ 线通过等利润线的最右点。

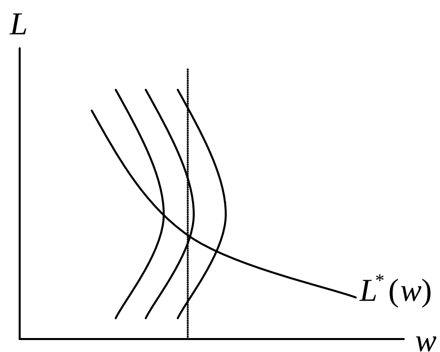


图.5.20　$L^*(w)$ 线通过等利润线的最右点

对工会来说，w 和 L 的增加都是好事，故其无差异曲线愈位于右上方就愈带来更高的效用或支付。

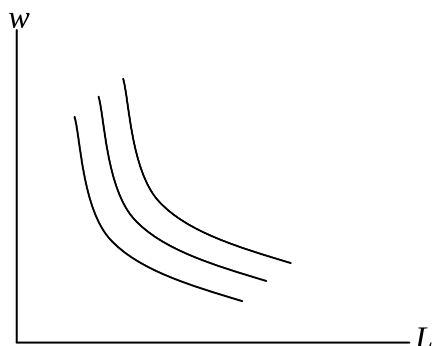


图 5.21　工会的无差异曲线

在第一阶段（倒数第二阶段），工会选 w^* 使 $U(w^*, L^*(w^*))$ 达到最大化，即工会在第一阶段的问题是

$$\max_{w\geq 0} U(w, L^*(w))$$

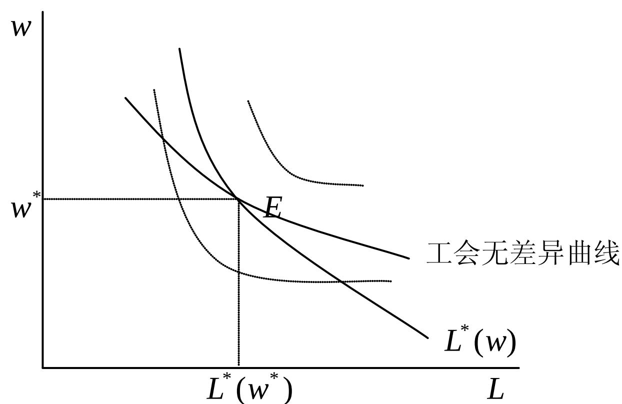


图 5.22　$L^*(w)$ 曲线与无差异曲线的切点决定均衡点

显然，$L^*(w)$与无差异曲线的切点 E 就是均衡点。

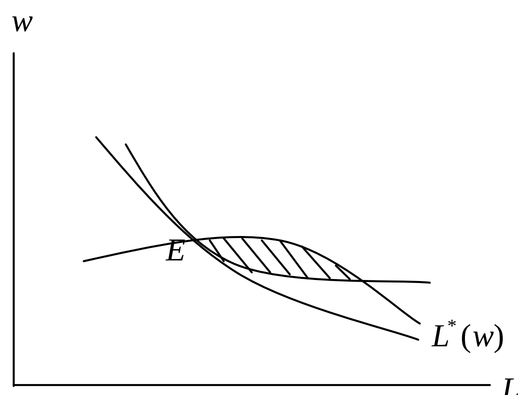


图 5.23　均衡的低效率

显然，若 w 与 L 的组合点位于图 5.23 中阴影部分，则工会与企业的支付都会同时提高，这说明在上述有限博弈中的均衡是低效率的。

Espinosa 与 Rhee（1989）基于如下事实为这一结果加以改进，即企业和工会之间经常会进行定期或不定期的重复谈判（在美国经常是每三年一次）。他们构造了一个无限次重复博弈模型，证明存在一个均衡，使得工会与企业选择的 w 和 L 位于图 5.23 阴影区域内，但在每一个一次性谈判中，这样的 w 和 L 都不是由逆向归纳法得到的解。

5.4.2 对“滞胀”的解释：宏观经济政策的动态一致性

对凯恩斯主义的一种精炼的概括是通过所谓“菲利普斯曲线”来表达的。增大就业率或减小失业率需要付出通货膨胀率上升的代价，而抑制通货膨胀又必然会引致失业率上升；换言之，低通胀与高就业率是对于政府来说是“鱼与熊掌不可兼得”的两难目标。这就是菲利普斯曲线（phillips curve）所表达的凯恩斯宏观经济学的基本教条。在 20 世纪 70 年代初，发达国家陷入了凯恩斯主义经济学未曾预料到的尴尬境地——高的失业率伴随着高的物价上涨率，这种现象被称为“滞胀”，即经济停滞与通货膨胀相伴而行。根据菲利普斯曲线，滞胀原本是不可能发生的，因为高失业率理应伴随低通胀率（甚至出现通缩），而高的通胀率也应与低失业率并存。作为凯恩斯主义经济理论必然推论的菲利普斯曲线现在不再成立意味着凯恩斯宏观经济学面临挑战，而 20 世纪 70 到 80 年代宏观经济学的复兴正是源于对传统凯恩斯宏观经济学提出众多挑战的结果。在对滞胀成因的解释中，不同的经济理论流派给出了不同的机理分析，其中芝加哥学派的分析可以说是影响最大的一类，以至于其主张被评论家们冠之以“货币主义学派”。尽管无论是弗里德曼还是卢卡斯们实际上都并未承认芝加哥大学有一个被称为“货币主义学派”的学术群体实际存在着。我们仍然按学术界的习惯以“货币主义”这一指称代表芝

加哥大学以弗里德曼为代表在对“滞胀”现象分析上的一类观点。

货币主义认为：滞胀实际上是政府长期推行凯恩斯主义宏观经济政策所导致的恶果。按照货币主义的看法，经济总是存在由外生（抑或内生？）因素(相对于凯恩斯主义政策工具而言)决定的正的失业率，这种失业率不可能通过需求管理所能改变，货币主义称其为“自然失业率”。但是，推行凯恩斯主义宏观经济政策的政府往往迫于选民的压力（或者好大喜功的冲动），总是试图通过扩张性的财政和货币政策进一步减低失业率。这种政策的直接后果是货币发行量增加速度过快（快于长期的经济增长率）和在短期有效的产出增加（就业上升）。短期效应鼓励政府推行这种政策，而有限期任期的政府又往往基于政绩表现的追求垂青于不断重复的短期繁荣而忽视了货币增加量增速过快带来的长期物价效应。短期出现的产出增长效应来自于货币增加后工人的工资刚性与物价上涨之间所带来的厂商实际工资成本下降效应。但一旦工会觉察到实际工资的下降就会有所行动迫使厂商提高工资率，实际工资成本在长期总会回升到政策变动之前的水平，于是厂商的最优化生产决策将导致产出和就业量回到政策变动之前的水平。因而在长期除了物价因政策变动而上升外什么也没有变。如果自然失业率因外生因素所致偏高的话，这种政策对经济停滞在长期是没有影响的。政府迫于选民压力或对政绩表现的追求不会无视经济停滞而无所作为，他们会再次推行新的一轮凯恩斯政策以干预经济，于是新一轮昙花一现的短期繁荣再度出现，接着又是回复到长期停滞的经济陷阱中去。凯恩斯政策的短期效应来自于菲利普斯曲线，而在长期菲利普斯曲线很难说是向下倾斜的，因而凯恩斯政策短期有效长期无效成为货币主义的一个基本观点。当政府不断推行上述扩张性的凯恩斯主义宏观经济政策时，其长期后果就是长期停滞依然而物价持续上涨，这就是货币主义对“滞胀”的基本机理解释。

为什么政府要沉迷于凯恩斯主义扩张性宏观经济政策带来的周期性经济兴衰的轮回之中呢？政府为什么不能遵循一种稳定的货币政策呢？货币主义及公共选择学派都认为政府基于对选票的追求对于即使是短期的繁荣也会情有独钟。但脱胎于货币主义的理性预期学派认为滞胀的现实中甚至消除了凯恩斯主义宏观经济政策的短期效应，因而凯恩斯主义宏观经济政策不仅长期无效，甚至连短期也是无效的。这是因为，公众不会在长期中屡犯同样的错误，一旦政府推出凯恩斯主义扩张性宏观经济政策，公众就预知物价会上涨，因而会立即要求上调工资以维持原有的实际工资水平，而在劳动市场上工会作为垄断力量有能力做到这一点。当然，在实际工资水平未发生任何变动的情况下，厂商不会对经济政策作出任何反应，因而产出和失业率依然不变，变化了的只是物价，包括工资率在内的所有物价都上升了，即作为一般物价水平的通胀率会出现明显上升。根据理性预期学派的这一说法，政府的扩张性政策实际上对就业未能产生一丁点的作用。但为何政府在预料到这种情况下会有政策冲动呢？关于这一点，我们将在这里给出的模型中表明，政府的政策冲动源于事后机会主义，即尽管政府可预知政策是无效的，

但政府在动态博弈中不出现政策冲动的事前承诺是不可置信的。给定公众相信政府不推行扩张性政策，政府的最优行动选择却正好是推行扩张性政策，而公众也预知政府的这一行为偏好，因而会形成正的事前预期通胀率，而这种预期将因公众对工资率的调整要求而实际推动物价上涨，从而在均衡时形成正的实际通胀率。因为公众对通胀率预期下的工资上调要求是为了维系不变的实际工资水平，因而产出及就业率不发生变化。货币主义及理性预期学派都认为失业问题如自然失业率过高应是凯恩斯主义政策以外的制度因素和自然因素才能影响的，仅凭凯恩斯主义政策难以解决。

由 Kydland 和 Prescott（1977）提出的货币政策动态博弈模型对上述思想给出一种具体的描述，以下介绍这一模型的基本情况。

例 5.9　货币政策的动态不一致性

模型中的局中人有两个，一是政府，其战略空间为在给定公众预期通胀率下所能选择的实际通胀率（货币政策）；二是公众，其战略空间为所选择的各种预期通胀率。

博弈行动顺序为：公众先动，政府在观察到公众的选择后行动。

支付函数：

政府同时关心通胀与失业问题（宏观经济政策的两大目标），故设政府的单阶段效用函数为

$$m(\pi, y) = c\pi^2 - (y - k\bar{y})^2,\quad c > 0,\ k > 1$$

其中 π 为通胀率，$\bar{y}$ 为自然失业率下的均衡产量，y 是实际产量。

$k > 1$ 的经济含义是由于市场扭曲（来自于工资刚性和市场的不完全竞争等）使自然失业率下的产量低于政府偏好的理想水平（即政府认为自然失业率过高），以及政府受到选民的压力而不得不寻求将产量提高到高于自然失业率产量的水平。该效用函数表明，尽管政府并不喜欢通胀，但若通胀能使产量提高到政府希望的水平 $k\bar{y}$，政府也会容忍某种程度的通胀。

产出与通胀之间的关系由含有通胀率预期的短期菲利普斯曲线决定。设定为

$$y = \bar{y} + \beta(\pi - \pi^e),\quad \beta > 0$$

其中 π^e 是公众预期的通胀率。

这种 phillips 曲线又称为“意外产出函数”，即只有未被公众预期到的通胀才会影响实际产出，其原因在于交易费用使得企业不可能随时调整工资率（以及工会的力量抵制工资向下调整）。

设政府在给定公众通胀预期下选择货币政策，则政府的优化决策为：

$$\max\ M(x, y) = -c\pi^2 - (y - k\bar{y})^2$$

$$s \cdot t \qquad y = \bar{y} + \beta(\pi - \pi^e)$$

将 $y = \bar{y} + \beta(\pi - \pi^e)$ 代入目标函数

$$M = -c\pi^2 - (\bar{y} + \beta\pi - \beta\pi^e - k\bar{y})^2$$

一阶条件

$$-2c\pi-2\beta(\bar{y}+\beta\pi-\beta\pi^e-k\bar{y})=0$$

$$(-2c-2\beta^2)\pi-2\beta[-\beta\pi^e+(1-k)\bar{y}]=0$$

$$\pi^*=\frac{\beta[\beta\pi^e+(k-1)\bar{y}]}{(c+\beta^2)}$$

π^*是政府短期最优通胀率（因它在计算π^*时视π^e为给定的，在长期中π^e会因π^*而变化调整，不是给定的）。

$(k-1)$可被理解为（政府认为的）扭曲程度。

上式表明：政府选择的通胀率是公众预期通胀率的函数，它就是政府的反应函数。

现假定公众有“理性预期”，则$\pi^e=\pi^*$。

代入反应函数，得到

$$\pi^e=\pi^*=\frac{\beta[\beta\pi^*+(k-1)\bar{y}]}{(c+\beta^2)}=\frac{\beta^2\pi^*}{c+\beta^2}+\frac{\beta(k-1)\bar{y}}{c+\beta^2}$$

故 $$\pi^e=\pi^*=\frac{\dfrac{\beta(k-1)\bar{y}}{c+\beta^2}}{1-\dfrac{\beta^2}{c+\beta^2}}=\frac{\beta(k-1)\bar{y}}{c}$$

图 5.24 表明理性预期$\pi^e=\pi^*$由反应函数$\pi^*(\pi^e)$与 45^0 线的交点决定。

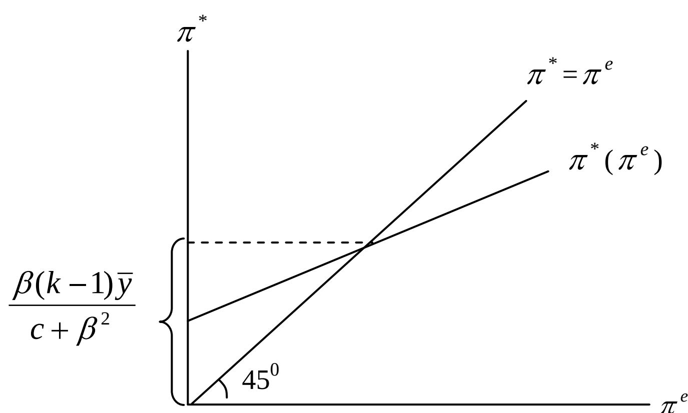


图 5.24 理性预期决定$\pi^e=\pi^*$

由$\pi^*=\pi^e=\dfrac{\beta(k-1)\bar{y}}{c}$知，$\beta$愈大（产出对未预料到的通胀率愈敏感），扭曲愈严重（$(k-1)$愈大），则理性预期通胀率（也是博弈均衡通胀率）就愈高。

当政府愈不喜欢通胀（c愈大时，$(-c)$是目标函数 M 中π的权数），均衡通胀率就愈低（与直观一致）。

此时，因政府选择的通胀率π^*被公众正确预期到（$\pi^*=\pi^e$），故实际产出水平将独

立于通胀（即 $y=\bar{y}$）。

政府一方面忍受着通胀之苦，另一方面又无法享受产出增加之益（减少失业带来选票的增加）。

将 π^* 代入效用函数并用 phillips 曲线消去 y，得到政府短期效用水平为

$$M_S=-c\left[\frac{\beta(k-1)\bar{y}}{c}\right]^2-\left[\bar{y}+\beta(\pi-\pi^e)-k\bar{y}\right]^2$$

$$=-(k-1)^2\bar{y}^2\left[1+\frac{\beta^2}{c^2}\right]$$

若政府选零通胀战略，则效用水平为

$$M_P=-c\cdot 0-\left[\bar{y}+\beta(0-0)-k\bar{y}\right]^2$$

$$=-(k-1)^2\bar{y}^2\text{(设公众也预测到零通胀率)}$$

这里下标 P 表示政府事前承诺零通胀率。

显然有 $M_S<M_P$，但为何政府不选择零通胀率呢？为何政府不一直按承诺的零通胀率行事呢？因为零通胀率不是可置信的承诺，即不是一个动态一致的政策，或者说不是一个子博弈精炼纳什均衡。

现假定政府许诺自己将实行零通胀率政策，且公众轻信了政府的许诺，则给定 π^e=0，政府的最优通胀率计算如下：

$$\max_{\pi} M=-c\pi^2-(y-k\bar{y})^2$$

$$s\cdot t\quad y=\bar{y}+\beta\pi$$

$$M=-c\pi^2-(\bar{y}+\beta\pi-k\bar{y})^2$$

一阶条件：$\dfrac{dM}{d\pi}=-2c\pi-2\beta(\bar{y}+\beta\pi-k\bar{y})^2=0$

$$\pi^*=\frac{\beta(k-1)\bar{y}}{c+\beta^2}$$

效用为 $M_f=-\dfrac{\left[(k-1)\bar{y}\right]^2}{1+\dfrac{\beta^2}{c}}$

其中 f 表示公众被政府愚弄的情形。

因 $M_f>M_P$，故政府无积极性兑现自己的许诺。

即给定公众相信通胀率为零，则政府一定会选大于零的通胀率。

因公众是理性的，且知政府是理性的（故预料到政府会如此干），故公众不会预期通胀率为政府所许诺的那样为零。因而有理性预期，效用只能为 M_S 而非 $M_P(>M_S)$。

这样，政府因无法使公众相信零通胀率而自受其若，即“聪明反被聪明误”。

货币主义的代表人物弗里德曼建议，可实行“单一的”货币政策，即以法律形式规定一个固定的货币增长率。注意，这是一个承诺行动，即将不可置信承诺的零通胀率变

为可置信的。因为这种法律规定限制了政府行动的自由（减小了政府的行动空间），它等价于一个可置信的承诺行动，但政府却反而受益（得到支付 M_P ）。

政府对自身声誉的考虑也是约束政府机会主义行动的一个重要因素（即考虑长期效用），因政府与公众之间的博弈实际上是一个重复博弈。如果有一个不制造通胀的强政府声誉，政府会长期受益的。

类似的例子还有：政府税收政策也存在动态不一致。如为了鼓励外资进入，有些发展中国家常许诺对外资的税收优惠政策。但在给定外资已进入情况下，政府又常常通过提高税率甚至没收外资企业的办法来增加财政收入。因此，除非有很好的法律制度保证政府言而有信，否则，投资者不会被政府的许诺所诱惑。

在这个例子中，若双方都预期和选择零通胀，则达到帕累托最优。但正如“囚徒困境”中的“（抵赖，抵赖）”一样，这并不是一个纳什均衡。当政府承诺零通胀时，若公众预期零通胀，则政府的零通胀政策就不是最优的了，不是一个子博弈精炼均衡。

5.4.3 世界石油产量与价格变动之谜的一种理论解释

①背景

在资源经济学中，有一个著名的 Hotelling 法则，根据该法则，当不考虑不可再生资源的生产（开采）成本时，资源价格在资源动态最优配置条件下，将以相等于市场利率的增长率而连续上升。假设 P_t 是第 t 期的资源价格， r 是市场利率，则 Hotelling 法则表明有如下关系：

$$\frac{P_{t+1}-P_t}{P_t}=r\text{或}P_{t+1}=P_t(1+r) \qquad (5.3)$$

在竞争性市场上，Hotelling 法则就是资源市场与资金市场之间达到均衡的条件（P. S. Dasgupta and G. M. Heal，1979）。杨秀苔和蒲勇健曾证明，Hotelling 法则也是动态一般均衡的必要条件，它刻画了不可再生资源动态配置中的最优行为（杨秀苔、蒲勇健，1993）。如果价格中包括了开采成本，则资源价格就是边际开采成本与资源租金率之和，而资源租金率仍然遵循 Hotelling 法则，即以市场利率作为增长率而不断上升（K. J. Arrow and S. S. L. Chang，1980）。

如果对资源的需求曲线不变，随着价格的上升，对资源的需求量就会不断下降，这样，资源产量在长期变动过程中就会呈现出一种不断缩减的趋势。在实际经济运行中，随着人口的增长、经济的发展，对资源的需求总是不断增加的，需求曲线也就发生相应的移动。多年来，理论家们几乎一致同意如下观点，即无论资源的价格和产量变动趋势会出现多少种复杂的组合，但有一点是肯定的，即资源价格持续下降与资源产量持续增加的现象是不可能同时出现的。按照资源经济学动态理论，资源产量持续增加说明需求上升，但需求上升又必然会导致价格上升而不是下降；当资源价格下降时，说明需求是下降的，但需求下降又必然导致产量缩减而不是增加。所以，建立在竞争性市场和动态

最优配置假定上的理论严格排除了不可再生资源价格下降与产量上升同时出现的可能性。

然而，当考察世界石油价格（牌价）与石油产量的实际数据时，人们发现，自 1960 年代以来，世界石油价格的长期变动趋势是持续下降的（除 1970 年代初因国际政治原因导致油价猛涨以外），而世界石油产量却几乎总是不断增加。这一事实令经济学家们大吃一惊，并感到迷惑不解。近十多年来，许多学者为了解开石油价格与产量变动之谜，进行了大量工作，并提出了一些模型，试图说明上述现象的发生机制。一些学者认为，石油价格下降与产量增加是由于石油工业出现持续的技术进步，使得开采成本持续降低。但是，这种理论迄今仍缺乏数据上的支持和具体的模型，不能令人满意。另外一些学者试图从资源存量的不确定性方面入手，通过运用不确定性最优控制数学工具，导出能解释上述现象的数学结果。K. J. Arrow 和 S. S. L. Chang 的工作集中体现了这样一种企图。他俩认为，Hotelling 法则成立的先决条件之一是资源初始存量是已知的，但实际上，石油储量总是随着世界范围内的勘探活动而不断变动，对于特定时期的定价者来说，未来储量的变动是不确定的。通过引入储量不确定假设，Arrow 和 Chang 导出了价格变动违反 Hotelling 法则的结果（K. J. Arrow and S. S. L. Chang，1980）。但是，他们的工作并未解释为何价格下降与产量上升会同时发生的现象，并且，在他们所给出的模型中，并不能导出价格持续下降的结果。迄今为止，经济学家们还未能对上述现象给出一个满意说明。杨秀苔和蒲勇健曾从一个新的角度，即从市场垄断和竞争的特殊假定，以及定价者追求既定的市场份额或尽量大的市场份额而不是长期（动态）利润最大化假定出发，给出一个寡头垄断的 Stackelberg 解，对世界石油价格与产量变动之谜给出一个满意的解释(杨秀苔和蒲勇健,1995)。

②世界石油价格与产量变动的特点

在 1950 年代后期，世界石油价格是上升的，但自 1960 年代初开始，世界石油价格便开始下降，在整个 1960 年代，石油价格一直是下降的。到了 1970 年代初，石油价格突然反弹，但如果按不变价格计算，1970 年代初的石油价格仍低于 1960 年代的石油价格。从 1970 年代中期以来直到现在，石油不变价格（即剔除通货膨胀因素）也一直呈平缓下降趋势，除了个别时期有间断反弹外，价格趋势是下降的，而产量的长期变动趋势是上升的。如何解释世界石油价格和产量的变动趋势呢？大多数学者认为，1950 年代后期的价格上升属正常现象，即稀缺性是价格上升的主要原因，这与 Hotelling 法则是一致的；而 70 年代初的价格反弹，应主要归于第三次中东战争的爆发，中东石油生产国为了对美国和西方进行制裁，通过 OPEC 把石油价格大幅度提高。另一些学者持不同看法，认为 1970 年代初的价格反弹主要是由于公众的预期产生了变化，由于公众（包括石油供给者）预期受到 1960 年代末和 1970 年代初以 Forrester 的“世界动力学”和 Meadows 的“增长极限论”为代表的资源危机论思潮的影响，认识到石油资源的有限性，

愿以较高的价格消费石油资源（P. S. Dasgupta and G. M. Heal，1979）。总之，可以认为 1970 年代初的反弹属于例外现象，把它放在经济模型的外生变量中去。人们关心的焦点是，如何解释自 1960 年代以来石油不变价格的长期下降和产量长期增加的趋势，这种违反 Hotelling 法则的现象导致资源动态配置明显偏离最优轨道。本文认为，石油价格上述变动特点的出现与 OPEC 在 1960 年代初期的形成密切相关。

在 20 世纪初期，一些受到美国地方政府协助的大型石油公司，开始到后来发现储有大量石油的中东和南美国家进行海外投资。这些公司被授权控制石油的生产和定价，而石油储有国仅获得固定的矿区使用费。因为石油储有国需要跨国公司的技术和设备，这种状况在 40 年代以前基本上维持不变。1940 年代以后，由于固定的矿区使用费不能抵消世界范围内的通货膨胀，以及储有国国内的民族主义兴起和大量新储量的发现，储有国开始试图改变原有的利益分摊方式。1943 年，委内瑞拉通过一项法律，后经 1945、1947 和 1948 年的数次修正，最终确定了原油生产利润必须在储有国与跨国公司之间进行分摊的新关系。这一法律的影响很快波及到中东国家。到了 1950 年，沙特阿拉伯协议被修正并纳入这一新关系。为了控制石油的生产销售，跨国公司和储有国规定一种石油牌价（Posted Price），该价格是垄断性决定的。尽管牌价可能会在长期发生变动，但与市场价相比要稳定得多。这种稳定性对于跨国公司来说可以减小激烈竞争的程度，对于储有国也有助于防止价格下降。本文中所提到的石油价格皆指不变的石油牌价。到了 1960 年，包括沙特阿拉伯、伊朗、委内瑞拉、利比亚和尼日利亚等十二个主要的石油生产国，为了完全控制世界石油生产和销售权，达成协议组成了石油生产销售的国际卡特尔 OPEC。自此以后，OPEC 成为世界石油生产中的最大寡头，在很大程度上左右了石油产量和价格。在 80 年代后期，OPEC 的石油产量占世界总量的三分之二，而在 60 年代初期，OPEC 的市场份额是低于这个数字的。

除 OPEC 外的石油生产国产量对 OPEC 的定价决策是有影响的，但非 OPEC 石油生产国中只有墨西哥有足够的储量对世界石油市场产生明显的影响。需要注意的是，墨西哥的储量是不确定的，还未精确探明（T. Tietenberg，1988）。

③OPEC 寡头垄断的斯塔克伯格博弈模型

根据上述分析，我们可以把世界石油市场的变动看成是 OPEC 与非 OPEC 两大集团之间相互作用的结果。OPEC 是一个有组织的卡特尔，即寡头之间由协议维系在一起的正式集团，而非 OPEC 只是一些石油产量较小的国家构成的一个集合，并非存在协议的真实集团。尽管如此，一方面由于非 OPEC 集团中有影响的主要是墨西哥，另一方面在模型上也可把全体非 OPEC 石油生产国视为一个集团，我们在以下的模型中，把世界石油产量和价格决定视为 OPEC 和非 OPEC 集团之间进行市场博弈的结果，并把所有石油生产者称为厂商。

1）厂商预期的需求曲线

由于与长期需求有关的变量是公开的，任何厂商都同样可能地了解这些变量，故假定所有厂商对世界石油需求曲线的预期是相同的。令X_t为第t期的世界石油需求量，Z_{it}为第t期世界经济对其它资源（包括人力资源和资本）的需求量，$(i=1,\cdots,n)$，y_t为第t期的世界经济总产出（如国民生产总值），则世界总量生产函数可写成

$$y_t = f(X_t, Z_{1t}, \cdots, Z_{Nt}, t) \tag{5.4}$$

两边取微分得：$dy_t = \frac{\partial f}{\partial X_t} dX_t + \sum_{i=1}^{N} \frac{\partial f}{\partial Z_{it}} dZ_{it} + \frac{\partial f}{\partial t} dt$

即
$$\xi_t = \frac{dy_t}{y_t} = \frac{\frac{\partial f}{\partial X_t} X_t}{y_t} \frac{dX_t}{X_t} + \sum_{i=1}^{N} \frac{\frac{\partial f}{\partial Z_{it}} Z_{it}}{y_t} \frac{dZ_{it}}{Z_{it}} + \frac{\frac{\partial f}{\partial t} dt}{y_t} \tag{5.5}$$

当$dt=1$时，ξ_t是世界经济年增长率，$\frac{dX_t}{X_t}$和$\frac{dZ_{it}}{Z_{it}}$分别是石油资源和其它资源需求量的年增长率，$\frac{\frac{\partial f}{\partial t}}{y_t}$是世界经济技术进步的年平均速度。此时，式（5.5）也可写成

$$\xi_t = \alpha \frac{dX_t}{X_t} + \sum_{i=1}^{N} \beta_i \frac{dZ_{it}}{Z_{it}} + A_t \tag{5.6}$$

其中，$\alpha = \frac{\frac{\partial f}{\partial X_t} X_t}{y_t}$，$\beta_i = \frac{\frac{\partial f}{\partial Z_{it}} Z_{it}}{y_{it}}$，$A_t = \frac{\frac{\partial f}{\partial t}}{y_t}$。如果按边际生产力计算要素报酬，$\alpha$和$\beta_i$分别是石油资源和其它资源所有者在世界总产出中的所得比例。在长期，厂商无法预测这些比例的变动方向，只能假定它们不变。假定厂商不能预测长期技术进步速度的变动方向，则A_t也只能被视为常数，记为$A_t = A$。式（4）可导成：

$$\xi_t - A = \alpha \frac{dX_t}{X_t} + \sum_{i=1}^{N} \beta_i \frac{dZ_{it}}{Z_{it}} \tag{5.7}$$

记$\eta_t = \xi_t - A$，它是纯粹由资源投入带来的经济增长率部分。厂商不能预测世界经济结构的长期变化方向，他们只能假定长期结构是不变的，即各种资源在生产中的比例搭配是不变的，即$X_t / Z_{it} = k_i$，$Z_{it} / Z_{jt} = k_{ij}$，或$X_t = k_i Z_{it}$，$Z_{it} = k_{ij} Z_{jt}$，k_i、k_{ij}为常数。由此得$\frac{dX_t}{X_t} = \frac{dZ_{it}}{Z_{it}}$，$\frac{dZ_{it}}{Z_{it}} = \frac{dZ_{jt}}{Z_{jt}}$，式（5）变为：$\eta_t = \left(\alpha + \sum_{i=1}^{N} \beta_i\right) \frac{dX_t}{X_t}$。

假定规模报酬不变（经济分析中常见的假定，在这里尤为可行），则$\alpha + \sum_{i=1}^{N} \beta_i = 1$，故$\eta_t = \frac{dX_t}{X_t}$。这说明：厂商对世界石油需求量的增长率（在任何价格上）的预期是依照他预期的世界经济增长率扣除预期技术进步速度来进行的。这里进一步（不失一般性）假定厂商的预期值η_t为常数$\eta_t = \eta$（不为常数的情形并不影响最终结果，但数学表达稍复杂一些）。尽管实际上石油需求曲线可能是非线性的，但厂商的预期可假定是线性的，

因为厂商通常使用线性回归技术（如最小二乘方法及其它线性计量经济模型）拟合出需求曲线，并作为其决策的依据。因此，可设预期的石油需求曲线为 $P_t = b_t - a_t X_t$，其中，$a_t > 0$，$b_t > 0$ 为常数，P_t 为价格。

在第 $t+1$ 期，对于任给价格 P，X_{t+1} 较 X_t 增加 $X_{t+1} - X_t = dX_t = \eta X_t$，即 $X_{t+1} = (1+\eta) X_t$。而在 P 下，原有 $P = P_t = b_t - a_t X_t$，在同一 P 下，有 $P = P_{t+1}$ 时，$X_{t+1} = (1+\eta) X_t$，或 $X_{t+1} = (1+\eta)\left(\frac{b_t}{a_t} - \frac{P_{t+1}}{a_t}\right)$，故有

$$\frac{a_t}{(1+\eta)} X_{t+1} = b_t - P_{t+1}$$

$$P_{t+1} = b_t - \frac{a_t}{(1+\eta)} X_{t+1}$$

由此可得到：$b_{t+1} = b_t$，$a_{t+1} = \frac{a_t}{(1+\eta)}$，或 $P_t = b_t - a_t X_t$，$b_t = b_{t-1} = \cdots = b_0$，$a_t = \frac{a_{t-1}}{(1+\eta)} = \cdots = \frac{a_0}{(1+\eta)^t}$

统一写为：$P_t = b_0 - \frac{a_0}{(1+\eta)^t} X_t$ （5.8）

式（5.8）就是厂商预期的需求曲线形式。

2）斯塔克伯格寡头垄断模型

将世界石油市场视为 OPEC 卡特尔与非 OPEC 石油生产国之间进行博弈的场所。经济学家已经为我们提供了许多分析寡头垄断的模型，但经过分析筛选，我们发现，只有斯塔克伯格（Stackelberg）份额解模型才能成功地解释石油价格和产量的变化特点。

设在 t 期，OPEC 的石油产量为 Q_t，非 OPEC 石油产量为 q，则 $X_t = Q_t + q_t$。斯塔克伯格模型假定，博弈双方中有一方占据着主动，他在采取任何对策时都要同时考虑对手的反应，并把这种反应纳入其最优对策选择的因素范围之中，这一方通常被称为领导者。在对策的另一方，由于其对市场的影响较小，在决策时不考虑自身的选择对对手的影响，只是在假定对手对策选定的基础上优化自己的利益，这一方通常被称为追随者(J. M. 亨德森，R. E. 匡特,1988)。这里，显然应假定 OPEC 是领导者，而所有非 OPEC 是追随者。由于最有影响的非 OPEC 石油生产国墨西哥的储量还未探明，故假定非 OPEC 石油生产国的储量未探明。再假定 OPEC 的目标是追求既定的市场份额或尽量使市场份额上升，而非 OPEC 石油生产国由于不可能在世界石油市场上占据可观的市场份额，其目标只能是追求即期利润极大化（因储量不明，不可能以长期利润最大化为目标）。

设 OPEC 追求的市场份额为 k。他们也十分愿意使 k 上升，但使 k 上升存在许多不确定性，因此，OPEC 的目标是追求市场份额不低于 k，但可大于 k。对于非 OPEC 集团，即期利润为 $E = P_t q_t$（不考虑开采成本的简化情形），极大化一阶条件为 $\dot{P}_t q_t + P_t = 0$，

字母上面的圆点表示对产量求导数。因 $P_t = b_0 - \frac{a_0}{(1+\eta)^t}(Q_t + q_t)$，追随者将 Q_t 视为常数，则有 $\dot{P}_t = -\frac{a_0}{(1+\eta)^t}$，故有

$$-\frac{a_0}{(1+\eta)^t} q_t + P_t = 0$$

或

$$-\frac{2a_0}{(1+\eta)^t} q_t + b_0 - \frac{a_0}{(1+\eta)^t} Q_t = 0$$

得

$$q_t = \frac{[b_0 - \frac{a_0}{(1+\eta)^t} Q_t](1+\eta)^t}{2a_0}$$

这就是非 OPEC 集团对 OPEC 产量 Q_t 的反应函数，OPEC 在决策时将考虑这一反应函数的影响，即将其代入需求曲线（5.8）：

$$P_t = b_0 - \frac{a_0}{(1+\eta)^t} Q_t - \frac{a_0}{(1+\eta)^t} \frac{[b_0 - \frac{a_0}{(1+\eta)^t} Q_t](1+\eta)^t}{2a_0} = \frac{b_0}{2} - \frac{a_0 Q_t}{2(1+\eta)^t} \quad (5.9)$$

OPEC 追求其市场份额 k，即 $\frac{Q_t}{(Q_t + q_t)} = k$，得：

$$(1-k)Q_t = kq_t = k\frac{[b_0 - \frac{a_0}{(1+\eta)^t} Q_t](1+\eta)^t}{2a_0} = \frac{kb_0(1+\eta)^t}{2a_0} - \frac{kQ_t}{2}$$

得：$Q = \frac{kb_0(1+\eta)^t}{a_0(2-k)}$

代入 $q_t = \frac{[b_0 - \frac{a_0}{(1+\eta)^t} \frac{kb_0(1+\eta)^t}{a_0(2-k)}](1+\eta)^t}{2a_0} = \frac{(b_0 - \frac{kb_0}{2-k})(1+\eta)^t}{2a_0}$

故 $X_t = Q_t + q_t = \frac{b_0(1+\eta)^t - \frac{kb_0(1-\eta)^t}{(2-k)}}{2a_0} + \frac{kb_0(1+\eta)^t}{a_0(2-k)} = \frac{b_0(1+\eta)^t}{a_0(2-k)}$ （5.10）

代入式（5.8）：

$$P_t = \frac{b_0(1-k)}{2-k} = b_0(1 - \frac{1}{2-k}) \quad (5.11)$$

显然，由式（5.10），X_t 是 t 和 k 的增函数；由式（5.11），P_t 是 k 的减函数。由此我们可以得到下列结论：

1）即使 OPEC 长期保持既定的市场份额，世界石油产量也是逐年上升的，只要厂商预期的 $\eta > 0$；但此时价格 P_t 不变。

2）如果 OPEC 在长期使其市场份额k逐渐上升，即$k = k(t)$且$\dot{k}(t) > 0$，则

$$\frac{dP_t}{dt} = \frac{-b_0\dot{k}(t)}{(2-k)^2} < 0$$

故价格P_t下降，但此时产量X_t显然仍然是上升的，这种情况很好地解释了本文开始所指出的石油价格与产量的长期变动特点。

事实上，OPEC 的实际决策过程并不是以长期利润最大化为目标，也不是以即期利润最大化为目标，而更多的是以占领世界石油市场为目标。OPEC 多年来就是要逐渐增大其在世界石油市场上的份额，并实际上也做到了这一点。当然，在个别的短期中，OPEC 可能会削减产量，从而使其市场份额下降，根据我们的模型，这会导致世界石油总产量和价格出现短期的波动（总产量不一定下降，因X_t随t指数上升可能抵消因k下降带来的产量下降，但价格会上升），这与实际情况很好地吻合。这样，我们发现，世界石油产量与价格变动的特有规律被很好地破译了，并由此可进一步得到如下结论。

3）厂商并未以长期利润最大化的动态优化原则定价，这直接导致资源配置的静态化和偏离动态优化轨道。

4）OPEC 可能在某时期并未使其市场份额上升到预期水平，即$k(t)$的增加$\Delta k(t)$小于预期的水平，但只要$k(t)$不下降，P_t仍下降（幅度小于预期水平），总产量X_t仍上升（幅度小于预期水平）。

5）厂商预期的η增大时，产量会加速上升。在 60 年代，由于世界经济的高速发展，η的预期值可认为较大，导致X_t增加得很快，这与实际情况十分吻合。

石油资源价格自 1960 年代以来呈现出长期下降和短期波动程度，而石油总产量基本上呈现出逐年增长（除个别时期以外）的现象，在这个模型中被成功地予以理论说明。这样，我们把石油价格和产量的长期变动趋势解释为 OPEC 与非 OPEC 石油生产国之间按领导者——追随者的 Stackelberg 寡头垄断博弈相互作用的结果，而它们的短期波动则被解释为纯粹的外生现象（是否如此，也许还需进一步深入研究）。

5.4.4 资产管理公司的最优股权

回购协议：一个博弈论模型

①导言

中国的国有企业在激烈的市场竞争中陷入经营性困境是近年来出现的一种普遍现象，导致这种现象发生的原因是多方面的。许多国有企业是在过去的计划经济条件下创建的，当时主要是基于国家经济发展战略的考虑而设置，随着市场经济体制的确立和发展，它们显得愈来愈难以适应市场规则，在产品结构、企业经营机制、公司治理结构和资本结构等方面显得相对落后和跟不上市场经济的发展，特别是沉重的债务负担使得国

有企业困难重重。在处于危境的国有企业中，有一些因特殊的原因而导致负债过重的企业本来有希望通过公司治理结构的优化、建立现代企业制度、技术创新和资本金注入而重振雄风，但由于在“七五”期间实行大规模技改所由国家划拨的大量资金在 1980 年代中期改由“拨改贷”形成大量债务（主要是利息）而难以轻装上阵，其发展陷入泥潭。这类企业通常是大中型的国有重点企业，工艺设备先进、规模庞大、产品有市场，但由于在过去的大规模技改期间只有投入而无产出（技改未完成期间的资金流量是负数），而“拨改贷”使其因利息累加而背上巨额债务，导致它们在近年代来因负债过重引起信用等级下降。另一方面，1990 年代后期国有商业银行开始实行资产负债比例管理，对贷款风险的意识增强，在企业再贷款中特别关注过去贷款余额的偿还情况，企业信用等级成了银行是否对企业进行贷款的重要参考参数。因此，在这种情况下，企业在完成技改后即使有潜在的获利能力，也难以在新的游戏规则下获得经营所需的新增贷款。同时，处于困境的企业往往又因不能偿还银行原有的贷款而使银行呆坏账增加，导致银行资产质量下降。这样一种银行——企业间的债务死结往往还会因恶性循环而拉紧，从而导致国有商业银行体系总体风险增加和企业危机加重的双败局面。为了解开这一死结、化解潜在的金融风险，政府作为国家对全体国有企业和国有商业银行的最高代理层，有权通过行政手段进行干预。于是，政府在 1999 年开始试行在 601 家国有企业进行债转股，成立四大家资产管理公司（信达、长城、华融和东方资产管理公司），将四大家国有商业银行和部分国有政策性银行在601家国有大中型企业中的大部分不良债权转为四家资产管理公司对企业的股权。具体的操作程序是：财政部注资成立资产管理公司；资产管理公司向国有商业银行出售由财政部担保的金融债券购买（以 1 比 1 的比例）银行不良资产；资产管理公司将买来的不良资产转为对企业的股权；资产管理公司成为企业股东（通常是大股东），参与企业重大决策，促进企业改制；资产管理公司进入阶段性持股，按市场化运作，其目标函数是在有限持股期内（现规定为 10 年）最大化地回收资产，基本思路是通过企业回购股权，向战略投资者、其它非国有企业、外资、私营企业甚至上市出售予社会投资者，最大化地收回原有资产，从而在改革过程中尽量减少国有资产的流失，做到国有资产保值增值，改革成本极小化。通过债转股，既可能化解国有商业银行体系中潜在的金融风险，也可能救活部分国有企业。这是因为，企业通过债转股减免了大量利息负担，提高了信用等级，同时还因债转股中完成了改制，从而轻装上阵，有可能引导银行加大资金支持力度，加快技术创新（包括新产品开发），最终实现扭亏为盈。在银行方面，由于债转股将不良资产转化为优良资产（财政部担保的金融债券），因而全面提高了资产质量，改善了银行资产负债状况，从而有能力加大贷款力度，同时，转股企业未来现金流的改善也降低了银行再贷款风险，使银行的贷款积极性增加，银行“惜贷”现象可望消除。这是一种预期的双赢结果。当然，金融风险只可转移不可能消除，它们被转移给了资产管理公司，最终实际上是转移给了国家财政部。作为政府部门

的财政部，其风险承受能力一般大于单一的商业性机构（银行或企业），因而这种转换过程有一定的合理性。

但是，导致许多转股企业处于困境的真正原因是多方面的，并不限于负债过重，如经营机制落后、产品结构不适应市场、技术创新能力差等。对于这类企业，希望通过债转股就一举扭亏为盈是不现实的，资产管理公司通过企业上市而出售股权的可能性也不大；同时，由于中国资产市场发育不足，资产评估中介业不发达，原有不良资产的真实价值难以准确评估且已远低于帐面价值，但现有的债转股思路要求资产管理公司不能通过打折出售股权，私营公司、外资及战略投资者也很难进入；所以，实际发生的情况多数将是通过转股企业的股权回购而实现资产管理公司的股权退出。据研究者的实际调研情况发现，现有的资产管理公司几乎都是通过与企业谈判，规定企业在未来若干年内逐年回购股权（通过以 1 比 1 比例）而设计其退出通道。

根据现有的资产管理公司经营机制，其经营的目标函数是资产回收数量，即财政部对资产管理公司的激励机制是根据资产管理公司每年回收的资产量而给予相应的奖励，并且这种奖励是即期的，可以假定资产管理公司每年经营的绩效收益与其当年回收的资产数额成正比。这样，资产管理公司的最大化行为是给企业规定最高最快的回购比例，即希望企业尽快尽量多地回购股权，但这样一来，企业的负担实际上可能没有减少多少，难以达到债转股的真实目的；另一方面，转股企业可能希望尽量少尽量慢地回购，因为企业通过将资金用于投资生产可加快企业发展和增强盈利能力。同时，这里面还有一个资金的时间价值问题。对于资产管理公司，由于其绩效收益与当年回收的资产量成正比，在贴现率大于零的情况下，它们总是愿意尽快地回收资产，因为未来的一块钱在今天不值一块钱，而回收资产总量是一定的，故宁愿回收今天的一块钱而不愿将这一块钱留待未来回收，再根据绩效收益与回收量的正比关系，资产管理公司宁愿获得今天的一元钱奖励而不愿将这一元钱奖励留待未来，所以，在现有的资产管理公司激励机制下，资产管理公司在与企业的回购协议谈判中，总是存在一种尽量要求企业回收股权的愿望甚至冲动。在另一方，企业宁愿将今天的一元钱留待未来按一元钱归还，这样既可以缓解企业今天的困难，待今后企业搞好了有充余的利润回购，又因未来一元钱在今天不值一元钱而获得资金的时间价值差额。所以，资产管理公司与转股企业在债转股谈判中通常就回购协议出现激烈的讨价还价过程，因为它们双方在回购的速度和数量上存在相反的偏好。研究者在实际调研中已大量观察到这种现象，但令人惊讶的是，通常这种协议最终还是达成了。对这种结果的一种解释是：尽管资产管理公司与企业存在相反的回购意向，但两者在谈判中的地位是不对称的，因为资产管理公司拥有是否将企业的债务转为股权的权力，而处于困境中的企业通常为了尽快减免利息负担，同时又因国有企业领导层的短期行为（因其任期有限，对短期政绩有强烈的表现欲），企业一般会被迫接受在它们看来是不合理和过于苛刻的由资产管理公司提出的回购协议条款，待转股减免了企业债

务负担后，企业的状况至少有一种次优改善，如果未来因过重的回购负担使企业再度陷入泥潭，企业预期政府会再次出面救助，这是国有企业固有的软约束使然。当然，可以认为这是一种可能的解释，但这种解释的不足是不能说明为何企业还要与资产管理公司进行激烈的讨价还价，也不能说明均衡是如何达成的。蒲勇健等曾经提出另一种解释，认为企业与资产管理公司之间就股权回购进行的讨价还价是一种博弈过程，并证明博弈的均衡点确定了协议规定的回购时序数量（蒲勇健，杨秀苔，2001）。他们通过构造一个简单的三周期完全信息动态博弈模型来说明此点。在这里，我们对此加以介绍。

②股权回购的完全信息动态博弈模型

1）基本假定

蒲勇健等采用 stackberg 动态博弈模型来刻画债转股中资产管理公司与企业之间就涉及股权回购的债转股协议谈判过程，这是一个讨价还价博弈，资产管理公司将提出股权回购的时序数额建议，企业选择同意或不同意，同时提出自己建议的股权回购时序数额。考虑到选择 stackberg 博弈模型，假定企业先动，资产管理公司在给定企业作出的最优战略选择下作出自己的最优战略选择。然后，企业再根据资产管理公司的反应函数确定最优的均衡选择，最后达到子博弈精炼纳什均衡。这里，首先对企业的战略空间作出说明。就企业来说，有权选择是否债转股，若已选择了债转股，则给定资产管理公司提出的股权回购时序量，企业选择自身努力程度，这种努力程度包括转股后企业自身的工作努力程度和进行经营体制转换即改制的努力程度。显然，转股后企业盈利或扭亏的能力与这种努力程度有正向变动的关系，不妨就用转股后企业的利润额来代表这种努力程度。因为企业在转股后愈是努力，包括愈是努力工作和进行改制，则在给定其他条件不变下，企业的利润就愈高（或亏损就愈小）。由于这里采用的是完全信息博弈，且假定资产管理公司至少能够回购部分股权（在现有的国有企业债转股思路中，预计相当多的企业不能在 10 年内完全回购股权，若资产管理公司在持股期内不能完全回收资产，则余额由财政部再出资补足），故不妨假定企业在转股后有逐年为正的利润流。根据蒲勇健等的调研，许多股权回购协议要求企业以利润、折旧基金甚至再借款来回购股权，为简化分析，这里假定企业只用利润流回购股权，更复杂的情况将在另外的论文中给出。其次，资产管理公司的战略空间是各种可选择的股权回购时序量建议。因资产管理公司的使命是债转股并最大化地回收资产，故不进行债转股的选择不在其战略空间内。下面给出模型的基本构架。

2）模型构架

局中人 1—企业；局中人 2—资产管理公司；

博弈的行动顺序如图 5.25。

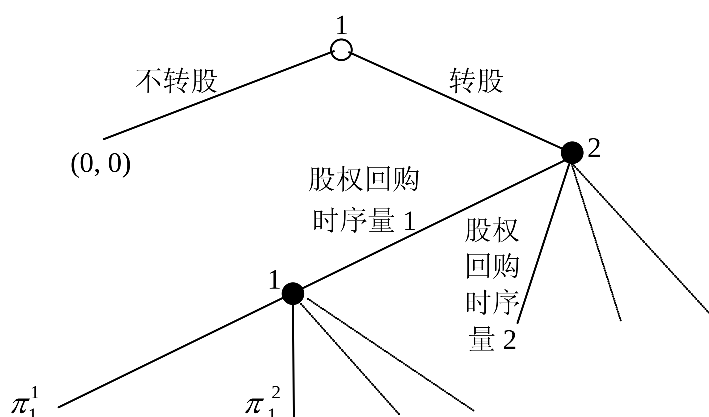


图 5.25 博弈的行动顺序

假设转股后企业的努力程度用转股后第一个生产经营周期（如一年）内的利润流π_1表示，并由假定随后各个生产经营周期内的企业努力程度不变（仍为π_1，但因随后的各个生产经营周期内的投资不同，故随后各生产经营周期内的利润将不同于努力程度）。在图 5.25 中，若企业选择不债转股，则企业的支付较低，用“0”表示，此时资产管理公司也偏离了其目标任务，故支付也较低，用“0”表示；若企业选择债转股，则资产管理公司可选择多种股权回购时序量，如图中的股权回购时序量 1，股权回购时序量 2……等等。若资产管理公司选择股权回购时序量 1，则企业可选择多种努力程度，如π_1^1，π_1^2……等等；在图 5.25 中，由于表达过于复杂，某些支付向量未标出。

我们用转股后企业在第一年中获得的利润π_1表示企业努力程度，并假定股权回购时序量（用回购资产数量表示）为$D_1, D_2 \cdots\cdots$，假设剩余$\pi_1 - D_1$用于企业在第二年的投资，则企业在第二年的利润增加额是$\pi_1 - D_1$的函数，记企业在第二年的利润为$\pi_2 = \pi_1 + f(\pi_1 - D_1)$，类似有$\pi_3 = \pi_2 + f(\pi_2 - D_2)$，$\pi_4 = \pi_3 + f(\pi_3 - D_3)$，……这里不考虑技术进步，故$f$与时间变量无关，假设$f(0) = 0,\ f' > 0,\ f'' < 0$，$\lim_{x \to 0} f'(x) = +\infty$（Inada 条件，该条件保证图 5.26（e）中的解是切点解）。实际上，在这种描述下，前述关于回购资金只来源于正的利润流的假设可取消。在回购资金可能来源于折旧基金或借款情形下，即$\pi_t - D_t < 0$时，只须假设$f(\pi_t - D_t) < 0$且仍单调增，此时仍能描述模型，利润流也不必假定为正。这是因为，挪用折旧基金和借款都会使企业未来盈利能力下降。故在未来利润的表达中出现减项。需要指出，尽管企业改制可在一个固定时期内完成，但因改制给企业内部人带来的内部人收益损失却是永久的，故企业内部人在转股后的每一年中都要损失一个固定的原有内部人收益，所以假定企业内部人在转股后的每一年中都要付出一个成本$C = C(\pi_1)$，总成本现值为$TC = \sum \delta_1^{t-1} C(\pi_1)$（若企业内部人关注无限期收益，则$TC = \frac{C(\pi_1)}{1 - \delta_1}$），假定导数$C'(\pi_1) > 0,\ C''(\pi_1) > 0$（边际成本递增），其中$\delta_1$为企业

内部人的贴现因子。我们这里假定企业在谈判中追求的是企业内部人（控制层）利益极大化。于是，局中人 1 的支付函数为 $u_1=\sum_{t=1}^{N}\delta_1^{t-1}(\pi_t-D_t)+\sum_{t=N+1}^{T}\delta_1^{t-1}\pi_t-TC$，其中 N 为资产公司持股期限（10 年），T 为企业内部人关注的收益年限（理论上可假设 $T=\infty$，当 T 取较小的有限值时，该模型也可用于分析企业行为短期化带来的后果）。

假定资产管理公司的支付函数为绩效收益最大化，并不妨设绩效收益与回收资产量成正比，则局中人 2 的支付函数为 $u_2=\sum_{t=1}^{N}\delta_2^{t-1}kD_t=k\sum_{t=1}^{N}\delta_2^{t-1}D_t$，其中 $k>0$，δ_2 为资产管理公司的贴现因子。约束条件：$0\le D_t\le\pi_t$，$\sum_{t=1}^{N}D_t\le D$，$\pi_t=\pi_{t-1}+f(\pi_{t-1}-D_{t-1}),t\ge 1$；其中 D 为转股资产总量。

局中人 1 的战略空间为 $S_1=\{(\text{转股, 不转股}),\pi_1,\pi_1\in[0,\infty)\}$；

局中人 2 的战略空间为 $S_2=\left\{(D_t),D_t\in[0,\pi_t],\sum_{t=1}^{N}D_t\le D,t=1,2,\cdots,N\right\}$

局中人 1 若在“转股”后选择了努力程度 π_1，则局中人 2 的最优选择满足：

$$\max_{\{D_t\}}u_2=k\sum_{t=1}^{N}\delta_2^{t-1}D_t,$$

$$s.t:\quad 0\le D_t\le\pi_t \qquad (5.12)$$

$$\sum_{t=1}^{N}D_t\le D,t=1,\cdots,N$$

设（5.12）的解为 $\{D_t^*\}$，则因

$$\begin{aligned}\pi_t&=\pi_{t-1}+f(\pi_{t-1}-D_{t-1})\\&=\pi_{t-2}+f(\pi_{t-2}-D_{t-2})+f(\pi_{t-2}+f(\pi_{t-2}-D_{t-2})-D_{t-1})\\&=\cdots\\&=\pi_1+f(\pi_1-D_1)+\cdots\\&=F(\pi_1,D_1,\cdots,D_{t-1})\end{aligned}$$

$$t=2,\cdots$$

故据式（5.12）必有 $D_t^*=D_t^*(\pi_1)$ （5.13）

式（5.13）给出了局中人 2 对于局中人 1 选择了 π_1 后的反应函数。

给定该反应函数，局中人 1 选择最优的 π_1 极大化自己的支付

$$\max_{\{\pi_1\}}u_1=\sum_{t=1}^{N}\delta_1^{t-1}\left(\pi_t-D_t^*\right)+\sum_{t=N+1}^{T}\delta_1^{t-1}\pi_t-\sum_{t=1}^{T}\delta_1^{t-1}C(\pi_1)$$

$$s.t:\quad \pi_1\ge 0 \qquad (5.14)$$

记（5.14）的解为 π_1^*，若 $u_1(\pi_1^*)>0$，则均衡为

$\{(\text{转股},\pi_1^*),\{D_t^*\},t=1,\cdots,N\}$

当 $u_1(\pi_1^*)<0$，均衡为

$\{(不转股,*),\{*\}\}$，其中*表示相应行动空间中的任意元素。

解此模型一般需要使用复杂的非线性动态规划方法，当模型中的时间变量取连续形式时，求解模型需采用最优控制理论中的哈密顿原理或庞特里亚金极大值原理法。由于篇幅所限，下面只给出一种三周期模型，可直接求解。

③一个三周期 stackberg 动态博弈模型

为了定性看出模型的含义，我们不妨将上述模型作一个极端的简化，令 N=2, T=3，此时，t 不再具有实际的年份含义，它只表示我们将持股期划分为二段（如每一段为 5 年），而 T=3 代表了持股期外的所有年份（包括无限期的情形），显然这是一个抽象模型，但它却能给出很直接的政策含义。为了得出解析解，我们将 $C(\pi_1)$ 的形式具体化，模仿信息经济学中通常的处理，设 $C(\pi_1)=\frac{1}{2}b\pi_1^2$，其中 $b>0$ 为常数（张维迎，1996）。

此时，

$$u_1=\pi_1-D_1+\delta_1(\pi_2-D_2)+\delta_1^2\pi_3-\frac{1}{2}b\pi_1^2(1+\delta_1+\delta_1^2)$$

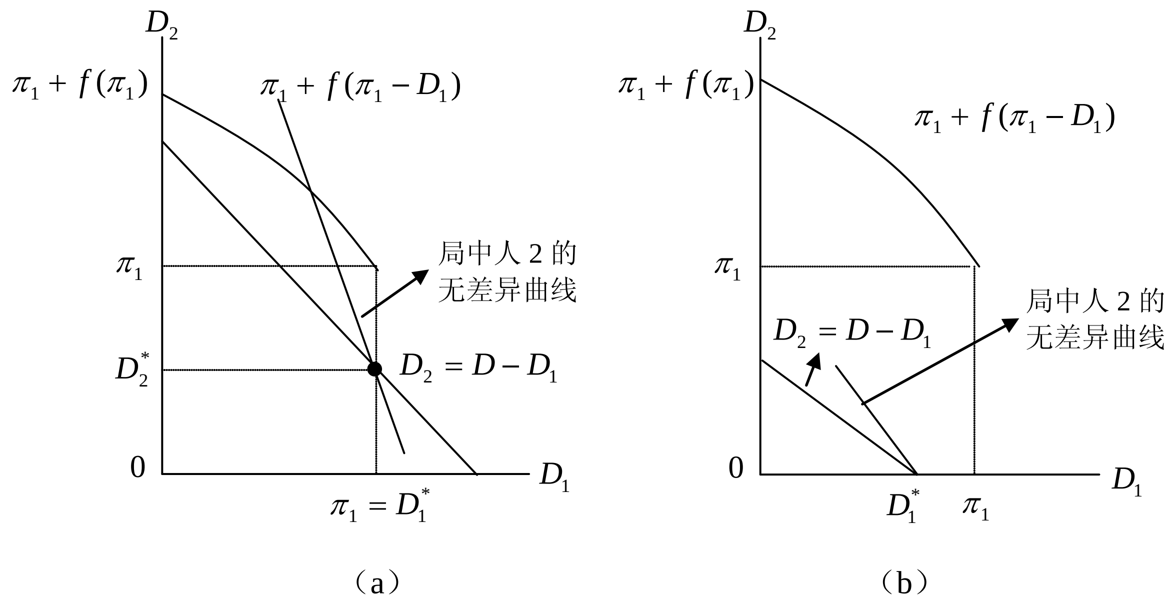
$$u_2=k(D_1+\delta_2 D_2) \tag{5.15}$$

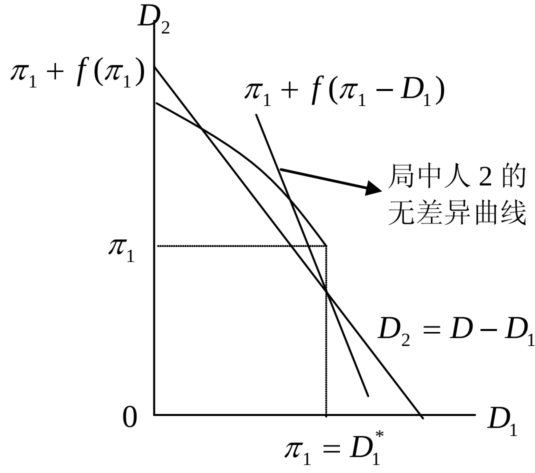
下面求解 $\{D_t^*\}, t=1,2$

$$\max k(D_1+\delta_2 D_2)$$

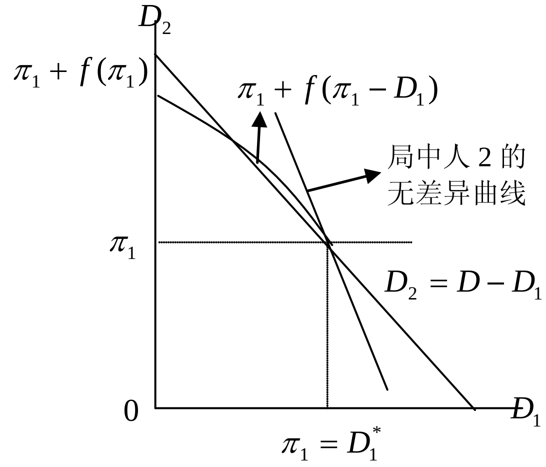
$$\begin{aligned} s.t.\quad & 0\le D_1\le\pi_1 \\ & 0\le D_2\le\pi_2=\pi_1+f(\pi_1-D_1) \\ & D_1+D_2\le D \end{aligned}$$

局中人 2 的无差异曲线为 $D_2=\frac{u_2}{k\delta_2}-\frac{1}{\delta_2}D_1$，因一般总有 $0<\delta_2<1$，故最优解有下述五种可能，见图 5.26。

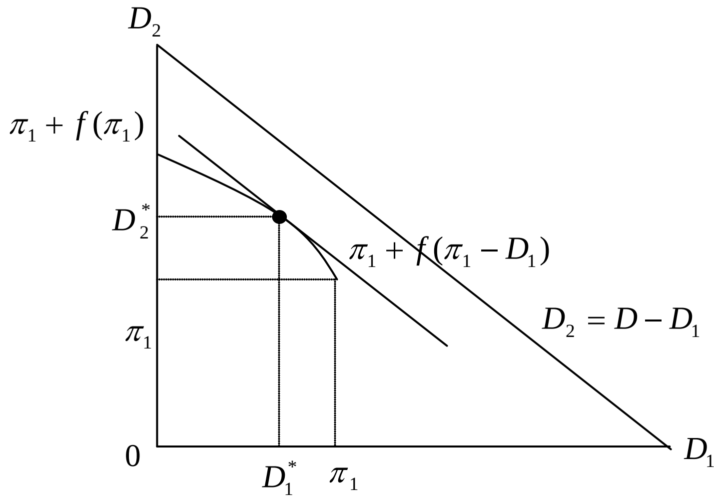




（c）



（d）



（e）

图 5.26　最优解的五种情形

情形（a）（b）（c）（d）都是角点解，为了避免数学上的复杂性，我们只考虑 D 充分大的情形，即情形（e），此时资产管理公司不能全部回收资产，其目标是最大化地回收资产。此时最优解是局中人 2 的无差异曲线与曲线 $D_2 = \pi_1 + f(\pi_1 - D_1)$ 的切点，故有

$$\frac{1}{\delta_2} = f'(\pi_1 - D_1)$$
$$D_2 = \pi_1 + f(\pi_1 - D_1) = \pi_2 \tag{5.16}$$

故 $\pi_1 - D_1 = A > 0$ 为常数，$\pi_2 - D_2 = 0$，由式（4）中第一式得：

$$u_1 = A + \delta_1^2(\pi_2) - \frac{1}{2}b\pi_1^2\left(1 + \delta_1 + \delta_1^2\right)$$

令 $\frac{du_1}{d\pi_1}=\delta_1^2-b\left(1+\delta_1+\delta_1^2\right)\pi_1=0$

$$\pi_1=\frac{\delta_1^2}{b\left(1+\delta_1+\delta_1^2\right)} \tag{5.17}$$

故最优解为：

$$D_1^*=\frac{\delta_1^2}{b\left(1+\delta_1+\delta_1^2\right)}-A$$

$$D_2^*=\frac{\delta_1^2}{b\left(1+\delta_1+\delta_1^2\right)}+f(A) \tag{5.18}$$

$$\pi_1^*=\frac{\delta_1^2}{b\left(1+\delta_1+\delta_1^2\right)}$$

$$u_1=A+\delta_1^2 f(A)+\frac{\delta_1^4}{2b\left(1+\delta_1+\delta_1^2\right)}>0$$

故均衡时局中人 1 会在第一个决策结上选择“债转股”，子博弈精炼纳什均衡为：

$$\left[\text{债转股},\left(D_1^*=\frac{\delta_1^2}{b\left(1+\delta_1+\delta_1^2\right)}-A,\ D_2^*=\frac{\delta_1^2}{b\left(1+\delta_1+\delta_1^2\right)}+f(A)\right),\ \pi_1^*=\frac{\delta_1^2}{b\left(1+\delta_1+\delta_1^2\right)}\right]$$

④政策含义分析

1）据式（5.16），资产管理公司的短期行为愈强，则δ_2愈小，故 A 愈小，由式（5.18）知D_1^*就愈大，D_2^*就愈小，此时资产管理公司就愈急于在持股前期回收资产；反之亦反。

2）据式（5.17），$\pi_1^*=\frac{1}{b\left(\frac{1}{\delta_1^2}+\frac{1}{\delta_1}+1\right)}$，故$\delta_1$愈大，$\pi_1^*$就愈大，因而企业愈是关注长远利益，其努力程度就愈高。

3）据式（5.18），b愈大，即企业内部人对债转股伴随的改制给他们带来的内部人利益损失评价愈高或企业努力的成本愈高，则企业的努力程度就愈低。

4）据式（5.18），企业努力程度愈高，资产管理公司提出的股权回购量就愈大，这是博弈论中常见的“棘轮效应”的一种具体表现。

5）据式（5.18），企业的努力程度与资产管理公司的类型（用δ_2刻画）无关，这一结论只限于我们所考虑的情形（e），但资产管理公司的股权回购时序量与自己的类型及企业的类型（用δ_1和 f 刻画）都有关。

在一般情形，我们应同时解出情形（a）、（b）、（c）、（d）和（e）的最优解，再比较各种情形中u_1的数值，取最大化u_1的解为局中人 1 的最优选择，然后得到子博弈纳什均衡，这种过程除了在数学上出现更多的复杂性外，基本结论是相似的。在情形（a）、（b）、（c）、（d）中，资产管理公司可在第一个周期内全部回收资产，这与我们观察到的情况不一致，事实上，绝大多数转股企业的状况难以保证做到这一点，这也是我们在

三周期模型中只考虑情形（e）的原因。

5.5 讨价还价模型与 Rubinstein 定理

博弈论通常运用于一些讨价还价博弈（bargaining game）的分析，而讨价还价是一类常见的现象。关于讨价还价博弈，大量的研究已使其成为博弈论的一个重要分支领域，本书在此不可能做到对其进行全面的介绍，而只是就轮流出价的讨价还价模型中及其相关的重要定理即 Rubinstein 定理作出说明。

Rubinstein 于 1982 年提出了轮流出价的讨价还价博弈模型（Rubinstein，1982），该模型属合作博弈模型。其简化的情形是假设有两个局中人 1 和 2 共同分配大小为 1 个单位的蛋糕，1 先动，提出分配方案，这称为 1 先“出价”；2 接着在观察到 1 提出的方案后选择接受或拒绝，如果拒绝，2 再提出自己的分配方案，称为 2 的“还价”，然后再由 1 考虑是否接受；若 1 接受，博弈就结束，否则 1 再出价，……，直到有一方的出价被另一方接受为止。这是一个完美信息动态博弈，见图 5.27 给出的博弈树。

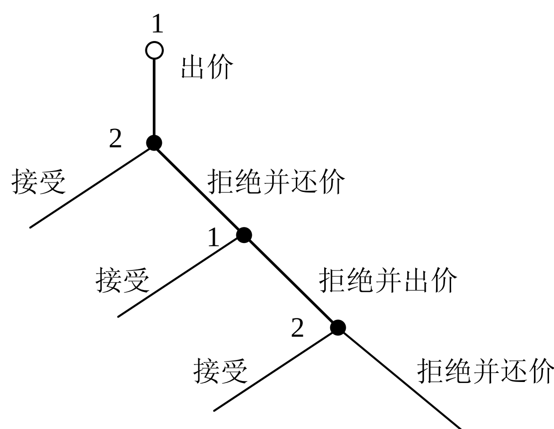


图 5.27　轮流出价的讨价还价博弈

令 x 表示 1 的份额，则 $(1-x)$ 就是 2 的份额，记 x_1 和 $(1-x_1)$ 分别是 1 出价时 1 和 2 的份额，x_2 和 $(1-x_2)$ 分别是 2 出价时 1 和 2 的份额，并设两个局中人的贴现因子分别为 δ_1 和 δ_2。于是，若博弈在时刻 t 结束，且在 t 时刻是局中人 i 的出价阶段，则局中人 1 和 2 的各阶段支付贴现值总和作为博弈支付函数就分别为 $\pi_1 = \delta_1^{t-1} x_i$ 和 $\pi_2 = \delta_2^{t-1}(1-x_i)$。当博弈是无限次进行下去时，博弈就成为无限次完美信息博弈，1 在时刻 1，3，5，……出价，2 在时刻 2，4，6，……出价。Rubinstein 证明，此时博弈有无限多个纳什均衡，但却只有一个“子博弈精炼纳什均衡”，该均衡由下面的 Rubinstein 定理给出。

5.5.1 有限次博弈的情形

当讨价还价博弈是有限次博弈时，我们可运用逆向归纳法求解其精炼均衡。

设博弈在 T 时刻结束。

若 T=2 时，则 2 最后行动，2 必提出 $x_2=0$，1 会接受（1 没有出价机会了，假定局中人在“接受”与“拒绝”之间无差异时，他选择接受）。

此时 2 的支付贴现值为 $\delta_2 \cdot 1=\delta_2$，故 1 在 t=1 时出价 $(1-x_1)\geq\delta_2$，2 会接受。

精炼均衡结果为：$x=x_1=1-\delta_2$，$1-x=\delta_2$

若 T=3，在 T=3 时由 1 出价，1 必选 $x_1=1$

在 $t=2$ 时，1 的支付贴现值为 $\delta_1 \cdot 1=\delta_1$，故 2 出价 $x_2=\delta_1$ 时 1 会接受，2 在 $t=2$ 得 $1-\delta_1$；

2 的支付在 $t=1$ 时的贴现值为 $\delta_2(1-\delta_1)$，故 1 在 $t=1$ 出价 $1-x_1=\delta_2(1-\delta_1)$，2 会接受，结果 1 得到 $[1-\delta_2(1-\delta_1)]$。

精炼均衡结果为：$x=1-\delta_2(1-\delta_1)$，$1-x=\delta_2(1-\delta_1)$，类似可得，T=4 时的精炼均衡结果为

$$\{[1-\delta_2(1-\delta_1(1-\delta_2))],[\delta_2(1-\delta_1(1-\delta_2))]\}$$

若 $\delta_1=\delta_2=0$（他们可能要忙于去干其它更重要的事），则精炼均衡结果总是 $(1,0)$，即两人都绝对无耐心时，则先出价者获全部蛋糕。

若 $\delta_2=0$，则无论 δ_1 如何，精炼均衡结果总为 $x=1$；但若 $\delta_1=0$，$\delta_2>0$，则精炼均衡结果为 $x=1-\delta_2$，即若 2 在 $t=1$ 拒绝了 1 的出价，则 2 在 $t=2$ 得到整个蛋糕，其支付贴现值为 δ_2，于是 2 在 $t=1$ 会接受任何 $1-x_1\geq\delta_2$ 的出价，即 $x_1\leq 1-\delta_2$，故 1 出价 $x_1=1-\delta_2$。

上述几种情形，均衡结果皆与 T 无关（设 $T\geq 2$）。现考虑另外情形：假定 $\delta_1=\delta_2=1$（双方都有无限耐心），则

若 T=1,3,5,……, 均衡结果为(1,0)

若 T=2,4,6,……, 均衡结果为(0,1)

即有“后动优势”，最后出价的局中人拒绝所有自己不能得到整个蛋糕的出价，一直等到博弈的最后阶段由自己出价获得整个蛋糕。

可以证明：若 $0<\delta_i<1$，$i=1,2$，则均衡结果与 $\frac{\delta_1}{\delta_2}$，$T$ 和谁在最后阶段出价有关。

但结果对 T 的依赖关系在 T 很大时变小。当 T 趋于无穷大时，若 $\delta_1=\delta_2=\delta$，则得到唯一的均衡结果。

$$\left(\frac{1}{1+\delta},\ 1-\frac{1}{1+\delta}\right)$$

因 $0<\delta<1$，故 $\frac{1}{1+\delta}>\frac{1}{2}$，即有“先动优势”。

5.5.2 无限次博弈与 Rubinstein 定理

当讨价还价博弈是无限次进行时，逆向归纳法不能直接使用，但我们可以运用逆向归纳法的思想以及博弈树在自身结构上的自相似性（即每一个子博弈在结构上相似于原博弈）解出其唯一的子博弈精炼均衡，这就是著名的 Rubinstein 定理。

定理 5.3（Rubinstein，1982）

若$T=\infty$，则轮流出价的讨价还价博弈有唯一的子博弈精炼纳什均衡，其均衡结果为

$$x^*=\frac{1-\delta_2}{1-\delta_1\delta_2}$$

当$\delta_1=\delta_2=\delta$时，$x^*=\dfrac{1}{1+\delta}$

证明：从图 5.18 可看出，从 1 出价的任何阶段开始的子博弈应等价于从$t=1$开始的原博弈，这一点曾由 Shaked 和 Sutton 严格加以了证明（Shaked 与 Sutton，1984）。于是，我们可应用逆向归纳法的逻辑去求解精炼均衡。

假定在$t\geq 3$由 1 出价且 1 能得到的最大份额为M（显然存在）。

1 在t得到的M对 1 来说等价于他在$t-1$得到$\delta_1 M$

故 2 在$t-1$出价$x_2\geq\delta_1 M$时，1 必接受，而 2 不会出比$\delta_1 M$更多的给 1，故 2 在$t-1$出价$x_2=\delta_1 M$对 2 是最优的，2 获得$1-\delta_1 M$。

在$t-2$，2 的最大支付贴现值为$\delta_2(1-\delta_1 M)$，1 在$t-2$出价$1-x_1\geq\delta_2(1-\delta_1 M)$时，2 会接受，而 1 不会出比此更多的给 2，故 1 出价$1-x_1=\delta_2(1-\delta_1 M)$，1 的最大获取为$1-\delta_2(1-\delta_1)$。

因为从$t-2$开始的博弈与从 t 开始的博弈完全相同，故 1 在$t-2$能得到的最大份额一定与其在 t 能得到的最大份额相同，所以

$$M=1-\delta_2(1-\delta_1 M)$$

得

$$M=\frac{1-\delta_2}{1-\delta_1\delta_2}$$

再设 1 在 t 能得到的最小份额为 m（显然存在），类似推理可得

$$m=\frac{1-\delta_2}{1-\delta_1\delta_2}=M$$

因总有$m\leq x\leq M$，而$m=M$，故必有$x=m=M=\dfrac{1-\delta_2}{1-\delta_1\delta_2}$。

证毕！

Rubinstein 定理的推论是：精炼均衡结果由局中人的贴现因子或耐心程度决定。

给定δ_2，当$\delta_1\to 1$时，$x^*\to 1$；给定δ_1，当$\delta_2\to 1$时，$x^*\to 0$。即有“耐心优势”，

有绝对耐心的人总可以通过拖延策略来使自己独吞蛋糕。

譬如，在农贸市场买菜时，退休老太太有充分多的时间去捕捉价格信息和与小贩讨价还价，她们有足够的耐心与小贩周旋，因而菜贩们一般不会在她们那里赚多少钱。

当$\delta_2 = 0$，1 得到整个蛋糕；但当$\delta_1 = 0$，2 不能得到整个蛋糕，除非$\delta_2 = 1$，即没有任何耐心的 1 也可得到一点份额，原因是除耐心优势外，博弈还有“先动优势”。

当$\delta_1 = \delta_2 = \delta < 1$，$x^* = \dfrac{1}{1+\delta} > \dfrac{1}{2}$，这是先动优势。

当$\delta_1 = \delta_2 = 1$，$x^* = \dfrac{0}{0}$，是[0,1]中的任何数，博弈有无限多个精炼均衡（$x^* = \dfrac{1}{2}$是聚点均衡）。

贴现率或贴现因子可理解为讨价还价中的一种成本，类似蛋糕随时间推延而不断缩小，每轮讨价还价的成本与剩余的蛋糕成正比。讨价还价的另一种成本是固定成本。如工会与企业的谈判拖延了工期，企业要承受推迟出售产品的利息损失（与产值成正比）和不能按期交货的违约罚款（一般是固定成本）两种损失。

固定成本的一种特殊形成是外部机会成本。外部机会愈好（机会成本愈高），谈判者就愈处于不利地位。

譬如，中国在 1999 年就加入 WTO 的谈判中就处于不利地位，因为这是 20 世纪最后的机会，若中国在这一次谈判中不能入世，则 WTO 在 21 世纪前五年不再吸收新的成员国并且要就 WTO 新规则进行修订，中国若不能参加 WTO 新规则的修订，则新规则对于今后加入 WTO 的中国可能是不利的。因此，中国在 1999 年的入世谈判中存在较高的拖延成本，δ很小，在谈判中不得不作出较大让步。另外，中国在多年的入世谈判中策略使用不太恰当，因为当时过多地渲染中国要加入 WTO 的热情反而使对手认识到中国的贴现因子较高，故而提出十分苛刻的条件。

由于讨价还价博弈属于合作博弈的范畴，我们还将在合作博弈内容部分再补充更多的这方面内容。

5.6 重复博弈与合作行为

博弈论中最为著名的模型大概是第二章中介绍的“囚徒困境”模型，它告诉我们一个局中人之间非合作的故事。按照博弈论中关于局中人为理性人的假定：“囚徒困境”中局中人之间的不合作似乎是必然的结果。在本书前面部分所给出的诸多例子中，似乎给人这样一种印象，即不合作好象是博弈中司空见惯的现象。但是，从另一方面看，现实生活中存在着诸多的合作现象，并且，对于经济学家来说，制度的形成及演化必定需要合作行为作为其基础。对于一个社会来说，基于合作的道德规范都是以合作为前提的。因此，博弈论能否与这些重要的经济社会现象相协调对于博弈论的预测能力来说是至关重要的。

到目前为止，博弈论专家认为博弈论不仅能解释不合作现象，而且也能解释合作现象。博弈论专家认为，合作现象可能来源于两种条件，一是博弈的无限次重复，二是信息不完全。

在“囚徒困境”中，如果一个小偷考虑到出狱后还要与另一位小偷打交道，则出卖同伙可能存在成本，当他足够高地估计这种成本时，他也许就会选择“抵赖”。当两个小偷都如此选择时，合作结果就出现了，而博弈的均衡结果就会是（抵赖，抵赖）而不是原有的（坦白，坦白）。博弈论专家认为，当局中人考虑到短期决策的长远利益时，其选择就可能倾向于合作。这样，当博弈可重复多次时，局中人就会考虑短期决策对其长远利益的影响，因而原有的仅考虑短期利益的均衡就可能变为注重长期利益的均衡。当“囚徒困境”中的小偷在出狱后还会再次联合作案时，他们就不得不考虑今次出卖同伙会在今后受到同伙惩罚的可能，因而其选择“坦白”的战略就增加了一个额外的成本。

这样，我们就进入了重复博弈的领域了。在前面所介绍的动态博弈模型都有一个共同的特点，即不同阶段开始的子博弈结构并不完全相同，这类动态博弈称为“序贯博弈”（Sequential game），而这里将引入的是同样结构的博弈重复多次，每一次称为“阶段博弈”，这类博弈就是“重复博弈”。

5.6.1 有限次重复博弈

重复博弈的定义如下：

定义 5.7 一个动态博弈是重复的，如果：

①前一阶段博弈不改变后一阶段博弈的结构；

②所有局中人都观察到了过去的博弈历史；

③局中人的总支付是所有阶段支付的贴现值之和。

在重复博弈中，每一个阶段博弈也可是一个动态博弈，如多次进行房地产开发博弈。

重复博弈可能使局中人不得不考虑这一阶段的行为对后面阶段博弈的影响，即注重声誉，这样，原来在单阶段博弈中不会出现的“合作”均衡（如囚徒困境中的(抵赖，抵赖)），在重复博弈中就可能作为均衡出现，而这正是研究重复博弈的意义所在。

重复博弈对局中人短期决策的影响程度还取决于博弈重复的次数，显然，当博弈重复次数较少时，长远利益不太重要，这种影响也就不太大，合作行为就不会出现。但是，令人感到惊讶的是，即使博弈可重复任意次，只要是有限次，合作行为仍然不会出现。这一次又是理性人假定在起作用。因为，最后阶段博弈的结果是可以预知的，并且必然是不合作的，因为对于最后阶段博弈来说，考虑长远利益已无意义。既然最后阶段博弈的结果是预先就明确知道了，之前的行为不会影响最后阶段博弈的结果，因而倒数第二阶段博弈也无须考虑长远利益，所以必然是不合作的。以此类推，我们知道从倒数第一、

第二、第三，……直到第一阶段的博弈都是不合作的。看来，要获取合作解，还必须假设博弈是无限次的。当然，这是就完全信息博弈来说的。对于有限次重复博弈来说，只有当信息是不完全的才能出现合作解，这一部分内容留到后面第七章中去介绍。这样，我们看到，影响重复博弈均衡是否是合作解的两个主要因素就是博弈重复次数及信息的完全性。重复次数会影响局中人在短期利益与长期利益之间的权衡。重复次数很多时，局中人可能为了长期利益而牺牲眼前短期利益。这是重复博弈分析给出的一个强有力的结果，它为现实中观察到的许多合作行为和社会规范提供了解释（而这是制度经济学研究的一个重要的方向）。另外，信息的完全性也很重要，当一个局中人的支付函数不为其他局中人所知时，他可能有积极性建立一个“好声誉”以换取长远利益。这也可解释为何有些本质上并不好的人在相当长时间内会干好事。本章仅讨论重复次数对均衡结果的影响，关于信息完全性的影响如“声誉模型”等将在第七章介绍。

关于有限次重复博弈，如果阶段博弈只有一个纳什均衡，则重复博弈也只有一个平凡的精炼纳什均衡，即阶段博弈纳什均衡的重复。这意味着在有限次合伙作案中，囚徒们每次被逮住都会如实招供的。下面再看一个例子，即连锁店悖论。

例 5.10 连锁店悖论

一个垄断的百货销售集团在 20 个地区都设有连锁店，另外一家公司打算在这 20 个地区也设立 20 个连锁店销售同类产品与前一公司竞争。假定原来的垄断者（在位者）在每个连锁店每年有 300 万元利润，而一旦新的公司设立的连锁店（进入者）进入，则由于进入者存在额外的进入成本，在位者每个连锁店的年利润下降到 50 万元而进入者年利润为 40 万元；当在位者对进入者进行打击，譬如说通过降价或折扣降低售价时，在位者每个连锁店的年利润降为零而进入者因存在较高的成本，其年利润会降到-10 万元，即每年亏损 10 万元。现在的博弈问题是：当进入者进入每一个地区时，在位者都有两种战略即打击或默许；而进入者在每一个地区也都有两种战略即进入还是不进入。表 5.2 给出了阶段博弈的战略式表述，而重复博弈就是两家公司在这 20 个地区不断重复的市场进入及阻挠博弈，假设进入者依次从第 1，第 2，……，到第 20 个地区进行进入决策博弈。

表 5.2　连锁店悖论

		在位者：默许	在位者：打击
进入者	进入	40, 50	-10,0
	不进入	0,300	0,300

根据划线法，我们知这个阶段博弈只有一个纳什均衡（进入，默许），并且不难验证它没有混合战略纳什均衡。

从直观上看，在第 1 个地区的博弈中，在位者为了使进入者不敢在别的地区开店，它会选择打击，但实际上这种威胁是不可置信的。我们用逆向归纳法的逻辑来分析：

在图 5.28 中，考虑第 20 个地区的博弈。因这是最后一个地区，打击对在位者无意义，其最优选择是“默许”，进入者进入。

再看第 19 个地区，因进入者和第 20 个地区上的博弈结果必是在位者默认，它进入，结果是确定的，不受这次博弈的影响，故知在位者必选“默许”，“打击”的威胁是不可置信的，故它必进入。

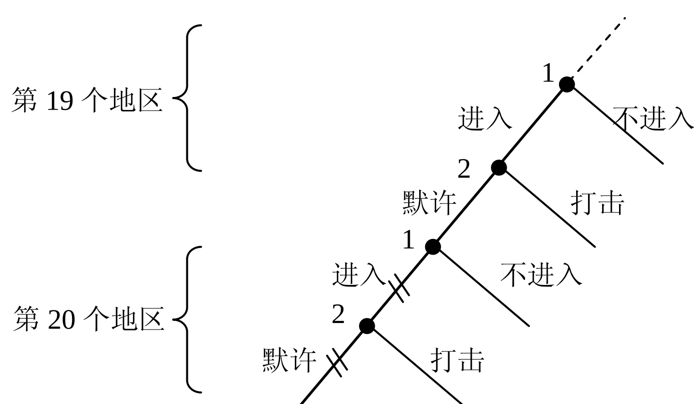


图 5.28　连锁店博弈

显然，如此倒推，知每一个地区的阶段博弈均衡必是（进入，默许），图 5.28 中的逆向归纳法求解表明这是唯一的子博弈精炼均衡。

由此，我们有如下定理。

定理 5.4　令 G 是阶段博弈，$G(T)$是 G 重复 T 次的有限次重复博弈$(T<\infty)$，则当 G 有唯一的子博弈精炼纳什均衡时，重复博弈 $G(T)$的唯一子博弈精炼纳什均衡结果是阶段博弈 G 的纳什均衡重复 T 次的结果。

局中人的总支付等于各阶段博弈的支付贴现值之和，而局中人的战略由局中人在各阶段博弈中的战略组成。因此，给定其他局中人的战略，也就给定了其他局中人在各阶段博弈中的战略。当其他局中人的战略由他在各阶段博弈中的子博弈精炼纳什均衡战略组成时，当且仅当该局中人的战略为各阶段博弈中的子博弈精炼均衡战略，该局中人的总支付达到最大。因该局中人是任一位局中人，因而我们就证明当单阶段博弈的子博弈精炼纳什均衡重复 T 次时，就构成整个重复博弈的一个纳什均衡。下面证明这个均衡是子博弈精炼均衡。

当一个子博弈是倒数第一个阶段博弈中的子博弈时，显然据定理条件知该均衡战略组合在此子博弈上也是纳什均衡。当一个子博弈开始于倒数第二个阶段博弈，给定其他局中人的战略及其在此子博弈上的限制，一个局中人的战略在此子博弈上的限制必为其最优战略。这是因为，对该局中人来说，任一其它此子博弈上的战略在倒数第一个阶段上限制带来的最后一个阶段博弈的支付不会大于原战略在倒数第一个阶段博弈上的限制带来的最后一个阶段博弈支付为大；同时，子博弈位于倒数第二个阶段博弈部分上的限制实际上构成倒数第二个阶段博弈上的一个子博弈，因为支付是各个阶段博弈支付的贴现和，故该战略在子博弈位于倒数第二个阶段博弈部分上的限制是子博弈在倒数第二个阶段博弈部分上的最优战略，因而该战略在子博弈上的限制是一个最优战略（因为支付是各阶段博弈支付之和）。因该局中人是任意的，故该战略组合也是子博弈上的纳什均衡。由此方法逐个倒推，我们可以证明该战略组合在重复博弈的任一子搏弈上的限制都是纳什均衡。所以该战略组合是一个子博弈精炼均衡。

最后再证明该重复博弈只有这一个子博弈精炼均衡。

记这个子博弈精炼均衡为 S，设有一个子博弈精炼均衡 S'，它与上述子博弈精炼均衡在最后一个阶段博弈上的所有子博弈上的限制都是纳什均衡，因而 S 和 S' 在最后一个阶段博弈上的限制是最后一个阶段博弈上的子博弈精炼均衡，由定理条件知它们在最后一个阶段博弈上的限制是完全相同的。在倒数第二个阶段博弈上，S' 在其上的所有子博弈上的限制必定是纳什均衡，故也是此阶段博弈的子博弈精炼均衡，由定理条件知 S 与 S' 在倒数第二个阶段博弈上的限制是完全相同的。由此倒推至第一个阶段博弈，可证 S 与 S' 在所有阶段博弈上的限制都相同，因而有 S 与 S' 相同。

证毕！

5.6.2 无限次重复博弈

在博弈是有限次进行的情况下，完全信息重复博弈不可能带来合作行为，因而我们将合作行为的出现寄望于无限次重复博弈。当然，在现实生活中，博弈应该说都是有限次的，但若局中人不知道博弈会在什么时候结束，每一次阶段博弈之后都有可能再一次博弈，此时局中人实际上认为博弈是可能进行无限次的。无限次重复博弈是对这种情况的一种分析方法。当然，如果是这样理解无限次重复博弈，似乎进入了不完全信息博弈的范围。但在博弈论中，不完全信息一般指局中人不知道其他局中人的支付函数，对于博弈进行的次数一般不假定存在不完全信息。当局中人不清楚博弈在未来什么时候结束（譬如，每个人都知道生命有限，自己总有一天会去世，但不知道将离开人世的确切时间，因而对每个明天都充满希望），但假设他知道未来每一个时刻博弈结束的概率。为简单计，假定博弈在每一个阶段结束的概率为 P，且各个阶段博弈结束的概率是相互独立的，则博弈到达 t 阶段的概率为 $(1-P)^{t-1}$，局中人在 t 阶段的期望支付为 $(1-P)^{t-1}\pi_t$，

其中π_t为局中人在t阶段的支付，设其贴现因子为δ，则局中人在t阶段的支付贴现值为$\delta^{t-1}(1-P)^{t-1}\pi_t$。令$\tilde{\delta}=\delta(1-P)$，则这个贴现值就可记为$\tilde{\delta}^{t-1}\pi_t$，总的支付为$\pi=\sum_{t=0}^{\infty}\tilde{\delta}^{t-1}\pi_t$。

这样，我们就可以将博弈视为无限次重复博弈的了。

首先，我们以“囚徒困境”为例子，说明在博弈进行无限次的情况下，合作行为就可能作为子博弈精炼均衡出现。

例 5.11　无限次囚徒博弈：走出囚境

我们根据表 2.2 给出的囚徒困境支付函数来分析当博弈重复无限次时的情形。假定两个小偷A和B的贴现因子为相同的常数δ，博弈重复无限次。我们将表明，当δ充分大时，合作均衡结果每阶段都为（抵赖，抵赖）将是一个子博弈精炼均衡。

对应于该均衡结果的均衡是什么呢？下面将证明，当两个局中人都选择所谓“冷酷战略”时，所得到的子博弈精炼均衡将生成这一结果。冷酷战略是指：

（1）在博弈开始选“抵赖”；

（2）若之前没有任何一方选“坦白”，就一直选“抵赖”，当之前至少有一方一旦“坦白”，之后永远选“坦白”（“冷酷”是任何一方一次不合作就触发了永远的报复(不合作)，这是必要的；因为若事后可以饶恕不合作行为的话，事前就会出现机会主义行为，合作行为就不会出现）。注意，在该战略中，一方自己一旦选了“坦白”，他之后也永远选“坦白”。下面证明两人都选冷酷战略时会构成一个精炼纳什均衡。先证它们构成一个纳什均衡。

给定囚徒B选冷酷战略，可证A选冷酷战略是最优的。

若之前没有人选“坦白”（包括在博弈开始时），A若选“坦白”，该阶段得 0 单位支付，但此举将触发B之后永远的报复，B会在之后永远选“坦白”，故之后A每阶段支付最多为－8，其总支付至多为

$$0+(-8)\delta+(-8)\delta^2+\cdots=\frac{-8\delta}{1+\delta}$$

而当A此时选“抵赖”并且之后每阶段都选“抵赖”时，B之后也会配合，每阶段都选“抵赖”，故此举使其总支付为

$$-1+(-1)\delta+(-1)\delta^2+\cdots=\frac{-1}{1+\delta}$$

当　$\frac{-8\delta}{1+\delta}\leq\frac{-1}{1+\delta}$

即$\delta\geq\frac{1}{8}$时，A选“抵赖”是最优的。

故当$\delta\geq\frac{1}{8}$时，A在没有人先选“坦白”时会选“抵赖”最优，并且，A之后每阶

段都选“抵赖”是最好的选择。

当 B 在之前已选了“坦白”，则 B 之后会永远选“坦白”，显然，给定 B 在此时及之后永远“坦白”，A 在此时选“坦白”是最优的，且 A 之后每阶段都选“坦白”是最优的。

当 A 在之前选过“坦白”，因 B 此时及之后必一直选“坦白”，故 A 在此时及之后一直选“坦白”是最优的。

所以，给定 B 选“冷酷战略”，A 选冷酷战略是最优的。由对称性知，当 A 选冷酷战略时，B 选冷酷战略也是最优的。因此，两人都选冷酷战略构成一个纳什均衡。下面证明该战略组合还是子博弈精炼均衡。

因为每个阶段博弈是一个静态博弈，因而每个子博弈的开始是从一个阶段博弈开始，且从任一阶段开始的子博弈与原博弈在结构上是相同的。

在冷酷战略纳什均衡中，可将子博弈分为两类，即：

类型 1：之前没有人曾经坦白；

类型 2：之前至少有一方曾坦白。

对于类型 1，实际上已经证明冷酷战略组合在子博弈上的限制是纳什均衡；

对于类型 2：冷酷战略在子博弈上的限制实际上是两人都一直选“坦白”，而显然这是一个纳什均衡。

由此证明：当 $\delta \geq \frac{1}{8}$ 即局中人有足够耐心或足够关注长远利益时，局中人都选冷酷战略构成无限次囚徒困境重复博弈的子博弈精炼纳什均衡，其均衡结果显然就是每阶段都为（抵赖，抵赖），它是帕累托最优的。这样，囚徒们走出了一次性博弈的困境[13]。

正如我们在第二章所述，“囚徒困境”实际上刻划了一系列社会经济现象中的不合作行为，因此，对于这些现象，如果将博弈变成无限次重复博弈，则在贴现因子充分大时也可望出现合作均衡。譬如，就古诺博弈来说，如果重复无限次，就会出现合作行为。

前述古诺博弈均衡产量 $q_1^c = q_2^c = \frac{a-c}{3}$，均衡利润 $\pi_1^c = \pi_2^c = \frac{(a-c)^2}{9}$，而完全垄断生产的产量为 $q^M = \frac{a-c}{2}$，垄断利润 $\pi^M = \frac{(a-c)^2}{4}$；当两企业只相遇一次时，则古诺博弈均衡是唯一的纳什均衡。但若博弈重复无限次，则某种形式的默契合谋就可能作为均衡结果出现。

冷酷战略为：

第 i 个企业首先选生产产量 $q^i = \frac{q^M}{2}$，继续选 $q^i = \frac{q^M}{2}$ 直到有一方选了 $q^j \neq \frac{q^M}{2}$，然后永远选 q_i^c，$i = 1,2$。

即从合作开始，若中途有任何一方不合作，则转入生产古诺均衡产量。

给定企业 j 选了冷酷战略，若企业 i 选合作，其每期利润为 $\frac{\pi^M}{2}=\frac{(a-c)^2}{8}$；若 i 选短期最优产量 $q_i=\frac{3}{8}(a-c)$，当期利润为 $\pi_i^d=\frac{9}{64}(a-c)^2>\frac{(a-c)^2}{8}$，但随后阶段利润流为 $\pi_i^c=\frac{(a-c)^2}{9}<\frac{(a-c)^2}{8}$。类似于囚徒困境无限次重复博弈中的证明，可得：

若 $\delta \geq \frac{9}{17}$，默契合作（合谋）就是一个精炼均衡结果[14][15]。

5.6.3 相关均衡与无名氏定理

在现实生活中，人与人或团体与团体之间一般不会是完全相互对立的或亲密无间的，更多的场合是介于两种极端关系之间。无名氏定理将要描述的正是这种情形。

"无名氏定理"（folk theorem）又被称为"一般可行性定理"（general feasibility theorem），它被 Friedman（1981）给出正式的证明之前就被人们广为所知。该定理的证明目前有许多不同的方法，这里介绍 Myerson（1991）给出的证明。不过，由于无名氏定理的证明需要引入一些相关的准备性概念，本小节将以适当的篇幅先介绍这些相关的概念。

一个十分重要的概念是"相关均衡"。什么是相关均衡呢？在我们已经知道的纳什均衡中，所谓均衡是指给定其他局中人所选取的战略，任何给定的局中人所选取的战略是最优的。当这个给定的局中人是任意的局中人时，战略组合就构成一个纳什均衡。这里，每个局中人在其自身给定的战略空间中自由地选择战略。如果我们将博弈规则作如下改变：在博弈中增加一个"信号"，每个局中人可以观察到信号。博弈规则规定，当局中人观察到信号后，他选择不同战略在给定其他局中人选择的战略下所得到的支付。这种博弈的纳什均衡就被称为"相关均衡"。

在这类博弈中，信号将起着在局中人之间进行通信或沟通的作用，这种作用会使得博弈均衡结果对局中人来说存在某种程度的帕累托改善。当然，引入信号的作用后，局中人的战略空间或许也会改变，因为局中人在参予信号的发送中也有选择，而这种选择就构成局中人新增的战略。

下面，我们首先通过一个例子来说明相关均衡的概念。

例 5.12　囚徒困境博弈的相关均衡

我们已知一次性囚徒困境博弈的纳什均衡不是帕累托最优的。表 5.3 给出一种在支付上与前面不一样的囚徒困境博弈矩阵，其纳什均衡为 (y_1, y_2)。

表 5.3　一种改变了支付的囚徒困境博弈

		2: x_2	2: y_2
1	x_1	2, 2	0, 6
1	y_1	6, 0	1, 1

现在，我们假定，在原博弈基础上再增加这样一条规则，即一个第三者向两位局中人出示一份契约，其上写道：“如果两个局中人都在这份契约上签名，则意味着每个局中人都许诺将选择 x_1 或 x_2 。如果这份契约只有一个人签字，则表示签名的局中人将选择 y_1 或 y_2 ”。每个局中人可以选择签名或不签名。这样，第 i 个局中人选择签名 s_i $(i=1,2)$ 就作为一个新增的战略进入其战略空间，战略空间因此扩大了。表 5.4 给出了这个相关博弈的支付矩阵。

表 5.4　囚徒困境的相关博弈

		2		
		x_2	y_2	s_2
1	x_1	2, 2	0, 6	0, 6
	y_1	6, 0	<u>1</u>, <u>1</u>	1, 1
	s_1	6, 0	1, 1	<u>2</u>, <u>2</u>

显然，这个相关博弈在原有博弈均衡基础上又增加了一个纳什均衡 (s_1, s_2) ，它的支付是原有博弈中的帕累托最优战略但它给出的支付为(2,2)。这里，契约是有约束力或强制性的，即局中人严格按照规则行动。当局中人 1 看见局中人 2 在契约上签了字，就意味着局中人 2 一定会按照规则进行战略选择，即若局中人 1 如果也签字，则局中人 2 一定会选择 x_2 ，而局中人 1 也相信局中人 2 一定会选 x_2 ，从而局中人 1 此时的最优选择是签字并选择 x_1 。反之，对于局中人 2 来说，情况也如此。但是，契约的这种强制性是表面上的，不是必需的假定。因为，在均衡位置给定一个局中人遵守契约规则，则另一个局中人的最优选择是遵守契约规则。当局中人 2 遵守规则，从而在看见两个人都签字后会选择 x_2 ，局中人 1 的最优选择是看见两人签字后遵守规则并选择 x_1 。相反，对局中人 2 也如此。这说明信号规则是“可实施的”。这一点对均衡位置是正确的，但对于非均衡位置就不一定如此了。当局中人 1 看见局中人 2 选择 x_2 时，如果他签了字但局中人 2 并未签字，则他的最优选择仍然是遵守规则并选 y_1 ；但反过来，当局中人 2 看见局中人 1 签了字但他自己未签字时，他的最优选择是选 x_2 但应签字，这与规则有出入。

在另一位置更加明显违反规则。如当局中人 2 不签字但选 y_2 ，局中人 1 签字这一位置。当局中人 1 看见仅自己签了字且局中人选 y_2 时，他的最优选择是遵守规则并选 y_1 ；但当局中人 2 看见仅局中人 1 签字时，给定他相信局中人 1 遵守规则，他的最优选择是选 x_2 且签字，不是不签字选 y_2 。因此，从均衡位置看，规则是可实施的，但从整个博弈看，要求契约要有强制力才行。

然而，两个人都签字的契约才是有效的，而当两个人都签字时，给定一方遵守契约规则，另一方的最优选择是遵守契约规则。反之，若至少有一方未签字，契约无效，博

弈实际上回到无信号的原博弈去了。所以，从这一点上看，契约本身无需作出强制性假定。

可以设想出更多的信号博弈规则，使均衡支付进一步获得帕累托改善。这是下面的例 5.13 给出的情形。

例 5.13　一个含有随机因素的囚徒困境相关博弈

现在，我们在上例中再加上第二份契约，它规定，如果两人都在这份契约上签了字，则掷一枚硬币，之后执行如下规则：若正面朝上则两人执行 (x_1, y_2)，若反面朝上则执行 (y_1, x_2)。这里，掷出的硬币带来随机性，正面朝上的可能性为 0.5。两人都签字时，每一个局中人获得的期望支付为（假定局中人是风险中性的）

$$0.5\times 0+0.5\times 6=3$$

则支付向量为(3,3)。

用 $\hat{s}_i$ 表示第 i 个局中人的在第二份契约上签字的选择，$i=1,2$。假定局中人只能在一份契约上签字，没有两个人同时签字的契约无效，且若只有一个人签字时此人只选择 y_1 或 y_2。表 5.5 给出了这一个博弈的支付矩阵。

表 5.5　加上第二份契约的囚徒困境相关博弈

		2			
		x_2	y_2	s_2	$\hat{s}_2$
1	x_1	2, 2	0, 6	0, 6	0, 6
	y_1	6, 0	$\underline{1}, \underline{1}$	1, 1	1, 1
	s_1	6, 0	1, 1	$\underline{2}, \underline{2}$	1,1
	$\hat{s}_1$	6, 0	1, 1	1, 1	$\underline{3}, \underline{3}$

这样，我们在此相关博弈中获得一个新的均衡 $(\hat{s}_1, \hat{s}_2)$，其支付(3,3)是前例中的均衡的帕累托改善。

下面，我们运用正式的数学语言为相关均衡进行定义。为此，先要给出一些准备性数学框架。

定义 5.8　给定任一战略式表述博弈 $G=\{S_1,\cdots,S_n,u_1,\cdots,u_n\}$，设局中人集合为 $\Gamma=\{1,\cdots,n\}$，某些局中人构成的集合 $g\subseteq\Gamma$，则 g 上的一个相关战略（Correlated Strategy）是指 g 中局中人所能选择的纯战略组合集 $S_g=\prod_{i\in g} S_i$ 上的任何一个概率分布 τ_g。

对于任一集合 S，我们今后用 $\Delta(S)$ 表示由 S 上所有概率分布密度组成的集合。

记 $\Delta(S_g)$ 为 s_g 上所有概率分布构成的集合，则相关战略就是 $\Delta(S_g)$ 中的元素。显然，$S=S_\Gamma$，$S_{-i}=S_{\Gamma/i}$，其中 $S=\prod_{i\in\Gamma} S_i$，$S_{-i}=\prod_{j\neq i} S_j$，局中人集 g 可以由一个调解人依照下

述方式随机地指定S_g中的纯战略组合来执行相关战略τ_g：对于S_g中任一$s_g=(s_i)_{i\in g}$，赋予s_g的概率是$\tau_g(s_g)$。由此，这个调解人会吩咐g中每个局中人i去执行战略s_i，即所指定的纯战略组合s_g的第i个分量。

给定所有局中人在ΔS中的任一相关战略u（即τ_Γ），对于每个局中人i，令$U_i(u)$表示当u在博弈G中被执行时局中人i所能得到的期望支付，即

$$U_i(u)=\sum_{s\in S}u(s)u_i(s)$$

再令$U(u)=(U_i(u))_{i\in\Gamma}$表示$\Gamma$中所有局中人由执行$u$所能得到的期望支付向量。

定义 5.9　$\prod_{g\subseteq\Gamma}\Delta(S_g)$中的任一向量$\tau=(\tau_g)_{g\subseteq\Gamma}$被称为一个契约或合同（contract）。对于任何一个这样的合同τ，如果g是在这个合同上签字的局中人集，那么τ_g就表示将被g中局中人实施的相关战略。因此，对于集

$$\{U(u)|u\in\Delta(S)\}$$

中的任一支付向量，都存在一个合同，使得如果所有局中人都在这个合同上签了字，那么他们将会得到这个期望支付向量。显然，所有可能的期望支付向量所组成的集是R^N的一个闭凸子集。

并非所有这样的合同都能被所有局中人实际签署的。譬如，在例 5.12 中的表 5.3 所示的博弈中，局中人 1 将拒绝在一份要求局中人实施(x_1,y_2)的合同上签字。如果局中人 2 签了字，局中人 1 签这样一个合同只能得到支付 0，但他可通过不签字并选择y_1而保证自己得到支付 1。对于任一局中人i，其在博弈G中的最小最大值（或安全水平（Security Level）是

$$v_i=\min_{\tau_{\Gamma|i}\in\Delta(S_{-i})}\max_{\tau_i\in\Delta(S_i)}\sum_{s\in S}\tau_{\Gamma|i}(s_{-i})\tau_i(s_i)u_i(s_{-i},s_i)$$

也就是说，局中人i的最小最大值是当所有其余局中人采用一个对i来讲最不利的相关战略来对付i时，他所能得到的最好期望支付。对付局中人i的一个最小最大战略（minimax strategy）是$\Delta(S_{-i})$中达到上式中最小值的相关战略。

上式可以改写为

$$v_i=\min_{\tau_{\Gamma|i}\in\Delta(S_{-i})}\max_{\tau_i\in\Delta(S_i)}\sum_{s_{-i}\in S_{-i}}\tau_{\Gamma|i}(s_{-i})\sum_{s_i\in S_i}\tau_i(s_i)u_i(s_{-i},s_i)$$

显然，最大化的τ_i和最小化的$\tau_{\Gamma|i}$是可以达到的，因为概率分布的极限仍是概率分布。当我们将上面的最大化和最小化理解为上确界和下确界时，它们都是可以达到的。

假设对任何i和s_{-i}，存在某$s_i(s_{-i})$使

$$s_i(s_{-i})=\max_{s_i^{'}\in S_i}\{u(s_{-i},s_i^{'})\}$$

此时，如果上式中的τ_i是最大化的，则必有

$$\tau_i(s_i^{'})=0，当u(s_{-i},s_i^{'})<u(s_{-i},s_i(s_{-i}))$$

故　$$\max_{\tau_i\in\Delta(S_i)}\sum_{s_{-i}\in S_{-i}}\tau_{\Gamma|i}(s_{-i})\sum_{s_i\in S_i}\tau_i(s_i)u_i(s_{-i},s_i)$$

$$= \max_{s_i \in S_i} \sum_{s_{-i} \in S_{-i}} \tau_{\Gamma|i}(s_{-i}) u_i(s_{-i}, s_i)$$

于是

$$v_i = \min_{\tau_{\Gamma|i} \in \Delta(S_{-i})} \max_{s_i \in S_i} \sum_{s_{-i} \in S_{-i}} \tau_{\Gamma|i}(s_{-i}) u_i(s_{-i}, s_i)$$

当且仅当

$$U_i(u) \geq v_i, \quad \forall i \in P$$

称相关战略 $u \in \Delta(S)$ 对于整个局中人 Γ 来说是“个人理性的”（individually rational），该不等式被称为相关战略 u 的“参与约束”（Participation Constraints）或个人理性约束（individual—rationality constraints）。

当合同是没有强制力的时，一个相关策略要得到实施必须是一个均衡。

考虑表 5.6 中的例子。这个博弈有 3 个纳什均衡：给出支付向量(5,1)的 (x_1, x_2)；给出支付向量(1,5)的 (y_1, y_2) 及给出期望支付向量(2.5, 2.5)的一个混合战略博弈均衡。

表 5.6　一个战略式博弈

		2: x_2	y_2
1	x_1	5, 1	0, 0
	y_1	4, 4	1, 5

其中 (y_1, x_2) 是一个帕累托最优的战略组合，但不是均衡。下面，假定在没有强制性合同情况下，通过“通信”来获取一个可带来帕累托改进的相关战略。

局中人抛一枚硬币，规定正面朝上则选择 (x_1, x_2)，反而朝上则选择 (y_1, y_2)。此时，给定一个局中人遵循规则，则另一个局中人遵守规则时的支付为

$$0.5 \times 5 + 0.5 \times 1 = 2.5$$

他不遵守规则时的支付必小于 2.5，因而尽管此规则对于局中人并不具有强制力，但局中人都会遵守此规则，这一规则是“自动实施的”（self-enforcing）。

在一个调解人（mediator，即一个能帮助局中人通信并共享信息的人或机器）的帮助下，这个博弈存在一个可自动实施的另一规则，它能够产生一个比上述(2.5, 2.5)更好的期望支付向量 $(\frac{10}{3}, \frac{10}{3})$。

假设此调解人将纯战略组合 (x_1, x_2)、(y_1, y_2) 和 (y_1, x_2) 中的每一个都以 $\frac{1}{3}$ 的概率随机地向两个局中人推荐，并假设每个局中人只知道他自己从调解人那里推荐来的纯战略，那么，尽管调解人的推荐不具有强制力，但这个借助于调解通信做变换后的博弈存在一个纳什均衡，其中局中人总是计划服从调解人的推荐。譬如，若局中人 1 从调解人

那里听到一个推荐选择“y_1”，则他想到局中人 2 被以相同的概率通知执行x_2或y_2，此时，他选x_1和y_1是无差异的。若局中人听到的推荐是“x_1”，则知道 2 被推荐的是“x_2”，此时他选x_1。因此，若局中人 1 预期局中人 2 服从推荐，他也会服从。同样的结论对局中人 2 也是成立的。此时的均衡支付向量为

$$\frac{1}{3}(5,1)+\frac{1}{3}(4,4)+\frac{1}{3}(1,5)=(\frac{10}{3},\frac{10}{3})$$

注意，实施这个没有合同的相关战略$\frac{1}{3}(x_1,x_2)+\frac{1}{3}(y_1,y_2)+\frac{1}{3}(y_1,x_2)$，要求每个局中人对于调解人的随机化结果只能收到彼此不同的局部信息。如果局中人 1 知道局中人 2 被通知选择x_2，那么，当y_1也被推荐给局中人 1 时，他是不会愿意选择y_1的。因此，在没有某种调解或噪音通信时，这个相关战略不可能被实施。

一般地，对于任一有限战略式表述博弈$G=\{S_1,\cdots,S_n;u_1,\cdots,u_n\}$，一个致力于帮助协调局中人行为的调解人至少需要让每个局中人i都知道推荐给他的是S_i中的哪个战略。但是，由于没有合同，局中人i在听到调解人的推荐后可以自由地在S_i中选择任何战略。因此，在具有调解通信的博弈中，每个局中人i实际上具有包括S_i到S_i的所有映射在内的一个扩大了的通信战略空间，其中每个通信战略都表示S_i中元素的一个可能规则，按照这个规则，局中人i选择S_i中的一个元素作为调解人从S_i中向他推荐的战略的一个函数。

假定调解人将按照$\Delta(S)$中的概率分布u决定其所要推荐的战略，这是局中人中间的共同知识；从而$u(s)$表示任一给定的纯战略组合$s=(s_1,\cdots,s_n)$会被调解人推荐给局中人的概率。于是，调解人所推荐并能为所有局中人遵循的战略是一个均衡的充分必要条件是

$$U_i(u)\geq\sum_{s\in S}u(c)u_i(s_{-i},\delta_i(s_i)),\forall i\in\Gamma\qquad\forall\delta_i:S_i\to S_i$$

此时称 u 是一个“相关均衡”。相关均衡是Γ中所有局中人在调解人的帮助下能自动地实施执行的相关战略，而调解人可以向每个局中人做出不具约束力的秘密推荐。

不难证明，相关均衡u满足下面的不等式方程组：

$$\sum_{s_{-i}\in S_{-i}}u(s)(u_i(s)-u_i(s_{-i},e_i))\geq 0,\ \forall i\in\Gamma,\forall s_i\in S,\forall e_i\in S_i$$

上式的含义是：给定任一s_i，若给定调解人所推荐的战略为s^*，则在预期所有其他局中人都服从调解人的这个推荐时，若将上式左边除以$\sum_{s_{-i}\in S_{-i}}u(s)$，就等于局中人$i$由服从调解人的推荐所得到的条件期望支付与其使用某个不服从该推荐的行动e_i所得到的条件期望支付之差。因此，上式断言，没有一个局中人i能够预期在得到调解人所推荐的任一战略s_i之后通过使用某个不服从的战略e_i来增大其期望支付。上述两个不等式组被称为“战略激励约束”（Strategic incentive constraints），因为它们表示了调解人所推荐的

相关战略必须满足的条件，只有满足了这个不等式条件，才有可能保证所有局中人都会理性地服从该调解人所推荐的S。

下面，我们给出重复博弈的一个一般性描述框架。

令Γ为一个非空的局中人集，又令Θ为一个非空集，表示可能的自然状态所组成的集。对于Γ中的每个局中人$i(i=1,\cdots,n)$，用非空集A_i和S_i分别表示局中人在博弈的每个阶段所能选择的行动（moves）集和所能收到的信号（signals）集。注意不要将这里的信号集S_i与前面用来表示战略空间的S_i相混淆。记

$$A=\prod_{i\in\Gamma}A_i,\quad S=\prod_{i\in\Gamma}S_i,$$

用q表示$\Delta(S\times\Theta)$中的一个初始分布，其中“$\times$”表示集合的笛卡儿集，与本书所用的记号“Π”相同，只不过后者用来表示多于两个以上的集合的笛卡儿集。P是转移函数（transition function）：

$A\times\Theta\to\Delta(S\times\Theta)$。局中人$i$的支付函数为$u_i:A\times\Theta\to R$。

记重复博弈为

$$G^r=\{\Gamma,\Theta,(A_i,S_i,u_i)_{i\in\Gamma},q,P\}$$

这种表述的重复博弈的解释如下：

这个博弈是以标号为 1，2，3，…的一个无限阶段序列进行的。在每个阶段（又称为“回合”），Θ中的某个状态是“目前的”（current）自然状态。每个局中人i关于自然状态的新信息是用其在阶段开始时所收到S_i中的一个信号来概括的。作为其在当前阶段和所有先前阶段所收到信号的一个函数，每个局中人i都必须在A_i中选择一个行动。θ^1为第一阶段的自然状态且每个局中人i在第一阶段开始时收到信号s_i^1的概率是$q(s^1,\theta^1)$。这里，记$s^k=(s_i^k)_{i\in\Gamma}$，$k=1,2,\cdots$。在任一阶段k，若目前的自然状态是θ^k，且局中人的行动组合是 A 中的$a^k=(a_i^k)_{i\in\Gamma}$，则$u_i(a^k,\theta^k)$是局中人i在阶段k中所得到的支付，而$P(s^{k+1},\theta^{k+1}/a^k,\theta^k)$是第$k+1$阶段的信号与自然状态将为$(s^{k+1},\theta^{k+1})$的条件概率。注意，这个概率是与第 1 至$k-1$阶段的状态与行动及第 1 至$k$阶段的信号都是无关的。

当且仅当存在某个正数Ω使

$$|u_i(a,\theta)|\le\Omega,\quad\forall i\in\Gamma,\quad\forall a\in A,\quad\forall\theta\in\Theta$$

时，称重复博弈G^r中的支付是有界的（bounded）。当且仅当

$$\sum_{s\in S}P(s,\theta|a,\theta)=1,\quad\forall_a\in A$$

时，称Θ中的状态θ是吸收的（absorbing）。当一个吸收的状态θ是第k阶段当前的自然状态，则θ将是k_i后所有阶段的自然状态。

假设每个局中人在每个阶段k都能记起其在第 1 至k阶段所获得的信号。为了保证他能记起他自己过去的行动，只需要在每个阶段令他的信号中有一个分量等于他自己在前一阶段中所使用的行动即可。因此，一般性的重复博弈G^r中局中人i的所有纯战略所

组成的集 C_i 是

$$C_i=\left\{c_i=(c_i^k)_{k=1}^{\infty}\middle|c_i^k:(S_i)^k\to A_i,\forall i\right\}$$

其中，$(S_i)^k$ 表示集 S_i 的 k 重笛卡尔积。类似地，局中人 i 的行为战略集 B_i 是

$$B_i=\left\{\sigma_i=(\sigma_i^k)_{k=1}^{\infty}\middle|\sigma_i^k:(S_i)^k\to\Delta(A_i),\forall k\right\}$$

即对于任一 k 和 $(S_i)^k$ 中任一信号历史 $(s_i^1,s_i^2,\cdots,s_i^k)$，一个行为战略 σ_i 必须确定 $\Delta(A_i)$ 中的一概率分布 $\sigma_i^k(\cdot|s_i^1,\cdots,s_i^k)$，它描述了在局中人 i 观察到这个信号序列下，i 将如何随机地决定其在阶段 k 的行动。令 $C=\prod_{i\in\Gamma}C_i$，$B=\prod_{i\in\Gamma}B_i$。

给定 B 中任一行为战略组合 $\sigma=(\sigma_i)_{i\in\Gamma}$，令 $P^k(a,\theta|\sigma)$ 表示这样一种概率，即若每个局中人在所有阶段都使用行为战略 σ_i，则在阶段 k，θ^k 将是当前的自然状态且局中人所选择的行动组合将是 d 的概率为 $P^k(d,\theta|\sigma)$。下面以归纳法方法具体地定义这一概率。

令 $\theta^1(\cdot|\sigma)$，$\theta^2(\cdot|\sigma)$，…和 $P^1(\cdot|\sigma)$，$\theta^2(\cdot|\sigma)$，……按如下归纳法方式定义：

$$\theta^1(s^1,\theta^1|\sigma)=q(s^1,\theta^1)$$

$$\theta^{k+1}(s^1,\cdots,s^{k+1},\theta^{k+1}|\sigma)=\sum_{\theta^k\in\Theta}\sum_{a\in A}\theta^k(s^1,\cdots,s^k,\theta^k|\sigma)$$

$$\cdot\left(\prod_{i\in\Gamma}\sigma_i^k(a_i|s_i^1,\cdots,s_i^k)\right)P(s^{k+1},\theta^{k+1}|a,\theta^k)$$

$$P^k(a,\theta^k|\sigma)=\sum_{(s^1,\cdots,s^k)}\theta^k(s^1,\cdots,s^k,\theta^k|\sigma)\prod_{i\in\Gamma}\sigma_i^k(a_i|s_i^1,\cdots,s_i^k)$$

如果每个局中人 i 都使用行为战略 δ_i，那么在阶段 k，$(s^1,\cdots,s^k)$ 将是各个局中人所收到的信号组合历史且 θ^k 将是当前自然状态的概率为 $\theta^k(s^1,\cdots,s^k,\theta^k|\sigma)$。

一个一般性重复博弈具有“完全状态信息（complete state information)，当且仅当在每个阶段，每个局中人都知道目前的自然状态。即对 Γ 中的每个局中人 i 和 S 中的每个信号 s_i，存在 Θ 中的某个状态 $w(s_i)$ 使得对于 Θ 中任意 θ 和 $\hat{\theta}$，A 中每个 a 及 $\prod_{j\in\Gamma\setminus i}S_j$ 中的每个 s_{-i}，均有

若 $\hat{\theta}\neq w(s_i)$，则 $P(s,\hat{\theta}|a,\theta)=0$

故除非目前的状态是 $w(s_i)$，否则信号 s_i 就决不可能发生（这里 $s=(s_{-i},s_i)$）。

在具有完全状态信息的重复博弈中，一个行为战略是“平稳的”（stationary)，当且仅当行动概率只依赖于当前的状态。也就是说，行为战略 σ_i 是平稳的当且仅当存在某个 $\tau_i:\Theta\to\Delta(A_i)$，使得对于每个 k 和 $(S_i)^k$ 中的每个信号历史 $(s_i^1,\cdots,s_i^k)$，都有

$$\sigma_i^k(\cdot|s_i^1,\cdots,s_i^k)=\tau_i(\cdot|w(s_i^k))$$

由于当上述方程成立时，τ_i 完全刻划了 σ_i，所以也可称任何这样的函数 $\tau_i:\Theta\to\Delta(A_i)$ 为局中人的一个平稳战略（stationary strategy)。这里 $\tau_i(a_i|\theta)$ 表示在每个阶段，当 θ 是自然状态时，局中人 i 将选择行动 a_i 的概率。

一个平稳战略组合 $\tau=(\tau_i)_{i\in\Gamma}$ 是一个均衡，当且仅当相应的行为战略组合（如上述

的 $(\sigma_i)_{i\in\Gamma}$）构成一个均衡。

一个具有标准信息的重复博弈（repeated game with standard information），或者一个标准重复博弈（standard repeated game，有时也称为超博弈(supergame)），就是一个其中只有一个可能的自然状态且局中人知道彼此过去行动的重复博弈。这就是说，在一个标准重复博弈中，$|\Theta|=1$，对每个 i 均有 $S_i=\prod_{j\in\Gamma}A_j$ 及

$$P(a,\cdots,a,\theta\,|\,a,\theta)=1,\ \ \forall a\in A$$

在具有标准信息的重复博弈中，一群人无限多次面临着完全相同的竞争局势，且对彼此过去的行为总是具有完全的信息。所以，标准信息结构使局中人相互间的反应能力达到最大。

标准重复博弈的平稳均衡恰好就是相应单阶段博弈均衡在每个阶段的重复。在一些弱假设下，几乎任何一个至少给每个局中人以其最小最大值的可行支付向量，都能在标准重复博弈的某个均衡中得以实现。对这一结论的一般陈述将在后面的无名氏定理给出并加以证明。在这里，我们通过对一个例子的讨论来揭示这一重要结论背后的基本直觉。

例 5.14　考虑由表 5.7 给出支付的重复斗鸡博弈。每个局中人 i 在每一阶段都有两个行动可供选择："谨慎"行动 a_i 和"大胆"行动 b_i。每个局中人在另一局中人选择谨慎时更喜欢选择大胆，但两个人都选择大胆时对双方来说都是最差的结果。帕累托最优的是两个人都选择谨慎。

表 5.7　重复斗鸡博弈

		2: a_2	2: b_2
1	a_1	4, 4	1, 6
	b_1	6, 1	-3, -3

如果局中人只玩一次此博弈，那这个博弈就有 3 个纳什均衡：一个给出支付向量(6,1)，另一个给出支付向量(1,6)，第三个给出期望支付向量的混合战略均衡(3,3)。该博弈还有一个相关均衡，其中 (a_1,a_2) 的概率为 1/2，且期望支付向量为(3.75, 3.75)；但在单阶段博弈的一个相关均衡中不可能达到更高的对称期望支付（即这是一个帕累托最优的相关均衡）。特别地，由于如果局中人 i 预期另一个局中人为谨慎（即选择 a_{-i}），则他将喜欢为大胆，所以 (a_1,a_2) 不是这个单阶段博弈的一个均衡。

现在假定这个博弈重复进行无限多次，具有标准信息，并且每个局中人都有相同的折现因子 δ，$0<\delta<1$。在这个重复博弈中，一个局中人的战略是一个规则，将他在每个阶段的行动决定为以前每个阶段所用行动的历史的一个函数。例如，像本例这样的重复博弈或重复囚徒困境博弈的一个著名战略即所谓针锋相对（tit-for-tat）。在这个战略下，

一个局中人在第一阶段选择谨慎行动，而此后则选择与其对手在前一个阶段所选择的行动相同的行动。

如果两个局中人都执行针锋相对战略，那么每个阶段的实际结果将是(a_1,a_2)，给每个局中人的平均支付为 4。这里，平均支付π_i按如下方式定义，当重复博弈在第k阶段的支付为π_i^k时，则

$$\pi_i=(1-\delta)\sum_{k=1}^{\infty}\delta^{k-1}\pi_i^k$$

这就是说，当有一个假想的重复博弈在每个阶段博弈的支付为相同的π_i时，则其总的贴现加权阶段支付和与原来的博弈是相等的。针锋相对不应与战略“总是选择a_i”混淆。在后一战略下，不论过去发生了什么，局中人在每个阶段都选择a_i。为了看出其究竟，首先假设局中人 1 执行“每个阶段都使用a_1”的战略，那么局中人 2 的最佳反应是一直使用b_2而得到平均支付 6。另一方面，如果局中人 1 执行针锋相对战略，且贴现因子接近于 1，那么由于局中人 2 因局中人 1 在下个阶段的报复行为所损失的要大于他在这一阶段选择b_2所得的，故局中人 2 决不会选择b_2。因此，一个战略中规定如何惩罚对手的非预期行为的成分可能是至关重要的，即使这种惩罚在均衡中不会被实施。

让我们检验说明，必须是多大的δ才能让针锋相对的战略制止对手的大胆行为。假定局中人 1 执行针锋相对战略，但局中人 2 思考是否在第一阶段大胆而后再回到谨慎。如果他这样做了，则他得到的支付在第一阶段是 6，第二阶段是 1，以后的每个阶段是 4，因而他的贴现平均支付为

$$(1-\delta)(6+1\cdot\delta+\sum_{k=3}^{\infty}4\delta^{k-1})=6-5\delta+3\delta^2$$

另一方面，如果他总是使用谨慎行动来对付局中人 1 的针锋相对战略，则他得到的贴现平均支付为 4。于是，为阻止局中人 2 的大胆行为，我们需要

$$4\geq(1-\delta)(6+1\cdot\delta+\sum_{k=3}^{\infty}4\delta^{k-1})$$

即$\delta\geq 2/3$。

尽管两个局中人都使用针锋相对战略是一个均衡，但这个均衡并不是一个子博弈精炼均衡。作为一个子博弈精炼均衡，局中人必须始终选择序贯理性的行动，即使在所给定战略下具有零概率的行动在历史上发生了，也理应如此。试考虑，如果局中人 2 在第 1 阶段偶然选择了b_2，局中人 1 将会面临怎样的局势。在两个人此后都执行针锋相对战略这一假定之下，每个偶数阶段的结果将是(b_1,a_2)，而每个奇数阶段的结果将是(a_1,b_2)；并且对局中人 1 的贴现平均支付是

$$(1-\delta)(1+6\delta+1\cdot\delta^2+6\delta^3+\cdots)=(1+6\delta)/(1+\delta)$$

另一方面，如果局中人 1 在阶段 2 选择a_1而偏离了针锋相对，且其后两个局中人都执行针锋相对，那么在阶段 1 之后，每个阶段的结果都将是(a_1,a_2)；因而，对局中人 1

的贴现平均支付是

$$(1-\delta)(1+4\delta+4\delta^2+4\delta^3+\cdots)=1+3\delta$$

不难证明：

$$当\delta>2/3时，1+3\delta>\frac{(1+6\delta)}{(1+\delta)}$$

再则，对于任一δ，由于支付 1 与 6 之间轮流选择要比总是得到-3 好一些，所以在两个局中人选择了(b_1,b_2)后，任何一个局中人都能从单方面偏离针锋相对战略而得到增益。因此，如果预期一个局中人此后也执行针锋相对战略，则另一个局中人执行针锋相对中的惩罚行动将是非理性的。即两人都实施针锋相对战略的局方案不是一个子博弈精炼均衡。

不过，存在一个修改针锋相对战略的方法可克服这个困难。考虑局中人i的下述战略：在每个阶段，除非对方过去选择b_{-i}的次数严格大于i选择b的次数，此时i会选择b_i之外，否则i将总是选择a_i。我们称这个战略为“扯平”（getting-even）战略。可以证明，只要$\delta>2/3$，两个局中人都使用扯平战略是一个子博弈精炼均衡。例如，如果局中人 2 在第 1 阶段偏离了这一战略而选择了b_2，但预计他此后会应用这一战略，那局中人 1 在第 2 阶段予以报复而选择b_1就是理性的。局中人 1 认为，在扯平战略下，局中人 2 会把他的报复b_1看作一个正当的反应；故局中人 2 应该不会在第 3 阶段回以反报复b_2。条件$\delta\geq 2/3$只是为了保证在两个过去大胆行动的次数相等时，没有一个局中人会去选择b_i。

由于针锋相对与扯平战略只是在犯了一个错误之后才有所不同，所以二者之间的差别很微小。令$r_i(k)$表示局中人i在阶段k_i前选择b_i的次数。如果局中人 1 正确地应用扯平战略，那么，不论局中人 2 使用什么战略，在任一阶段k，$r_2(k)-r_1(k)$将总是等于 0 或 1，且等于 1 的充分必要条件是局中人 2 在上一阶段选择了b_2时。因此，根据扯平战略，当且仅当局中人 2 在上一阶段选择了b_2时，局中人 1 才会选择b_1。这两个战略之间的差别只有在局中人 1 自己偶然偏离了其战略但随后又回过来实施那个战略的情况下，才变得明显。

如果$\delta<2/3$，则针锋相对战略与扯平战略都不能作为此博弈的一个均衡得到持续。然而，只要$\delta\geq 0.4$，此博弈仍有其它能在每个阶段都以概率 1 实现结果(a_1,a_2)的子博弈精炼均衡。

在假设两个局中人都在均衡中选择a_i时，局中人 1 能够阻止局中人 2 选择b_2的唯一条件是，他具有某种报复或惩罚的可置信威胁，一旦实施这一威胁，局中人 2 将来要付出比当前阶段从 4 增加到 6 更大的代价。不过，由于随着贴现因子愈来愈小，以后支付的损失较之于当前支付的增益也就愈来愈无关紧要，所以就愈来愈难以阻止局中人 2 向b_2的偏离。针对最低可能的贴现因子，为设计一个阻止这种偏离的战略，我们需要找出对局中人 2 来说代价最大的可置信威胁。

在最坏的情况下，局中人 2 总可以通过在每一阶段都选择a_2来保证自己在每一阶段都能得到支付 1。因此，局中人 2 在为偏离而受到惩罚时所得到的贴现平均支付不可能低于 1。如果在局中人 2 第一次选择b_2之后，局中人 1 永远选择b_1而局中人 2 永远选择a_2，这样的一种支付将确实可以实现。而且，局中人 1 永远选择b_1和局中人 2 永远选择a_2也是理性的，只要每个局中人都预期对方会这样做。因此，对局中人 2 最差的可置信惩罚将是局中人 1 永远选择b_1而局中人 2 永远选择a_2作为反应。

考虑局中人i的下述冷酷战略（grim strategy）：如果有任一过去的阶段，其中一个局中人大胆而另一个局中人谨慎（即结果为(b_1,a_2)或(a_1,b_2)），则在第一个其中i是谨慎局中人的这种阶段，那i在当前阶段及以后都选择b_i；否则i就选择a_i。也就是说，在冷酷战略之下，i一直合作（即“谨慎”）到对手背叛去选择大胆，在这种情况下，i通过永远大胆去惩罚对手。只要$\delta \geq 0.4$，两个局中人都执行冷酷战略是一个子博弈精炼均衡。

这种无穷的不宽恕的惩罚看起来相当极端。所以，我们或许有兴趣去寻找其它的战略，像针锋相对战略那样，能在贴现因子小于$2/3$时，在子博弈精炼均衡中，只有一个阶段的惩罚，而以后又能永远维持在(a_1,a_2)。确实存在一种这样的战略，我们称之为“相互惩罚”（mutual punishment）。对于局中人i，这个战略可被描述如下：

如果是在第一阶段，或上一阶段的行动是(a_1,a_2)或(b_1,b_2)，则i选择a_i；如果上一阶段的行动是(b_1,a_2)或(a_1,b_2)，则i选择b_i。如果两个局中人都被认为在执行这一战略，但一个局中人在某个特定的阶段偏离了这个战略，则在下一阶段，局中人应选择(b_1,b_2)，而这对他们双方都是最差的可能结果（“相互惩罚”由此得名）。如果局中人 2 在正确地执行这一战略，而局中人 1 偏离去选择了b_1，那局中人 2 将在下一阶段选择b_2作为报复，并将如此持续到局中人 1 再次选择b_1。其思想是，局中人 2 惩罚局中人 1 的偏离直至局中人 1 也参与对自己的惩罚之中，或在鞭子之下要露出胸膛。

为了检验两个局中人都执行相互惩罚战略是否是一个子博弈精炼均衡，我们必须考虑两种可能的偏离：在战略需要a_i时选择了b_i和在战略需要b_i时选择了a_i。如果战略需要a_i但局中人i在某个阶段偏离了，而另一个局中人执行相互惩罚战略，则i所得到的支付在这个阶段是 6，下一个阶段是-3；若他不偏离，则两个阶段都得到支付 4。由于下一阶段的支付要被一个附加因子δ贴现，所以若$4+4\delta \geq 6+(-3)\delta$，即$\delta \geq 2/7$，则这种偏离会被制止。另一方面，如果战略需要$b_i$但局中人$i$在某个阶段偏离了，而另一个局中人执行相互惩罚战略，则i所得到的支付在这个战略是 1，下阶段是-3；若他不偏离，则他在这个阶段所得到的支付是-3，下阶段是 4。因此，若$-3+4\delta \geq 1+(-3)\delta$，即$\delta \geq 4/7$，则这个偏离会被制止。这样一来，只要贴现因子至少是$4/7$，两个局中人都执行相互惩罚战略就是一个子博弈精炼均衡。

到目前为止，我们只讨论了这样的均衡，其实际结果总是在每个阶段都是合作的(a_1, a_2)，除非有某个人犯了错误而偏离了这个均衡。然而，这个重复博弈存在许多其它的均衡。原来一个阶段博弈的任何一个均衡（在重复时）都可能成为重复博弈的一个均衡。例如，存在一个局中人 1 总是选择b_1而局中人 2 总是选择a_2的均衡。这是一个对局中人 1 来说最好但对局中人 2 来说最坏的一个均衡。

这个重复博弈还有一些福利性质很差的均衡，并且对两个局中人来说都接近于最坏的情形。而且，其中的某些均衡又具有自然的或逻辑的吸引力，使得它们在某些情形下会成为人们实际实施的焦点。为了看出导致这些坏均衡的逻辑，注意到扯平均衡与冷酷均衡两者都具有这样的性质：以前愈胆大的局中人以后应该愈胆小。有些人或许认为相反的原理更合乎逻辑：过去愈胆大的局中人应该被认为是将来更胆大的局中人。

因而，考虑一个我们称i为$q-$阵地（$q-$positional）的战略，对每个局中人i来说此战略可定义如下：若i已选择b_i的次数严格大于另一局中人已选择b_{-i}的次数，i就选择b_i；若另一局中人已选择b_{-i}的次数严格大于i曾选择b_i的次数，i就选择a_i；而如果b_i和b_{-i}曾被选择的次数相同，i现在就以概率q选择b_i并以$(1-q)$的概率选择a_i。对这个战略的直觉阐释就是，已确立较强的大胆声誉的局中人在将来可能是大胆的，而已有较多谨慎行为模式的局中人应该保持其所创造的谨慎形象。当没有一个局中人有更为大胆的声誉时，则他们可以独立地以某种方式在大胆与谨慎之间随机化。

给定 0 与 1 之间贴现因子δ的任意一个值，存在一个q使得两个局中人都执行$q-$阵地战略是一个子博弈精炼均衡。为了计算这个q，首先注意到，尽管在这个重复博弈中并不存在内在定义的备择自然状态，但局中人的阵地战略可用隐性状态的术语来定义，这个隐性状态就是在过去的回合中，局中人 1 选择了b_1的次数与局中人 2 选择b_2的次数之差。在任一阶段，若隐性状态为正，且两个局中人都执行这个$q-$阵地战略，则从现在起每个阶段给出的支付都将是(6,1)。类似地，若隐性状态为负，则从现在起，每个阶段给出的支付都将是(1,6)。当隐性状态在第 1 阶段为 0（当然，正如实际上那样），且两个局中人都应用q阵地战略时，令z表示每个局中人预期得到的期望贴现平均支付（根据对称性，对两个局中人是相同的）。那么，当隐性状态为 0 时，局中人 1 乐意在于b_1与a_1之间随机化的充分必要条件是

$$
\begin{aligned}
&q((1-\delta)(-3)+\delta z)+(1-q)((1-\delta)(6)+\delta(6)) \\
&= q((1-\delta)(1)+\delta(1))+(1-q)((1-\delta)(4)+\delta z) \\
&= z
\end{aligned}
$$

这些等式表明，在局中人 2 是大胆的概率为q时，局中人 1 在第 1 阶段不论是大胆的（第一个表达式）还是谨慎的（第二个表达式），他在截断的博弈中所得到的期望$\delta-$贴现平均支付都一样的，且事实上等于z（正如我们定义z时所要求的那样）。给定δ，由这两个方程可解得q和z，结果如表 5.8 所示。因此，假如局中人有着长期目标，故δ

接近于 1，那么，阵地均衡给局中人的期望贴现平均支付只略好于局中人在最差的均衡中也能保证自己得到的最低支付。实际上，建立一个大胆型声誉的动机把两个局中人推向了接近于双方最差可能均衡的消耗战（war of attrition）之中。

表 5.8　对 4 个贴现因子的贴现平均支付和初始的大胆概率

δ	q	z
0.99	0.992063	1.0024
0.90	0.925	1.024
0.667	0.767	1.273
0.40	0.582	1.903

这里所讨论的各种均衡，可被看作是那些能在长期关系中形成的行为类型的例子。当具有长期关系的人们察觉到他们当前的行为将对彼此未来的行为产生影响时，他们可能会根据当前行为与未来行为之间所预期的联系形式，而理性地变得更合作（如在扯平均衡中）或更好战（如在阵地均衡中）。定性地讲，更合作的均衡看起来涉及到彼此之间的互惠联系（例如，“你今日之所为，正是我明日之所行”），而更好战的均衡则涉及到个人行为的外推联系（例如，“我今天这么做，明天我还会这么做”）。

我们从上述例子中所领会到的一般直觉是，在标准重复博弈中，当局中人有充分耐心时，几乎任何一个至少给每个局中人其最小最大值（即安全水平）的可行支付向量都可以在此重复博弈的某个均衡中实现。所谓可行支付向量指单阶段博弈中所有混合博弈所能达到的支付向量，包括非均衡的支付向量。这就是说，就一个标准重复博弈而言，其均衡中可行的支付向量一般可能会与相应具有合同的单阶段博弈中的可行支付相一致。事实上，表达这一直觉的一般可行性定理已在一系列弱技术假设之下得以叙述并证明。这些可行性定理被称为无名氏定理，这是因为，某些弱可行性定理，在任何严格的陈述发表之前，作为口头习俗的一部分，已被许多博弈理论家所知道或相信。

Rubinstein（1979）在超过准则下对标准重复博弈的子博弈精炼均衡证明了一个无名氏定理。Fudenberg 与 Maskin（1986）在贴现平均支付准则下对标准重复博弈的子博弈精炼均衡证明了一个无名氏定理。这里，所谓超过准则和贴现平均支付准则都是表达无穷次重复博弈支付函数的两种不同方法。本书使用的表述无穷次重复博弈的支付函数方法是单阶段支付的贴现加权和准则，它等价于贴现平均支付准则，因为两种准则表达的支付函数之间只相差一个系数$(1-\delta)$。用相关均衡代替纳什均衡，我们在这里叙述并证明 Fudenberg 与 Maskin 结论的一种表达形式。

为了加强这个结论，这里我们假设存在一个调解人或相关装置，该调解人或相关装置在每个阶段只是公开宣布所有局中人共同观察到的信息。

这个博弈的一个相关战略是任 $-u=(u^k)_{k=1}^{\infty}$，使得对于每个 k，u^k 是从 $(A)^{2k-2}$ 到 $\Delta(A)$ 的一个函数。这里 $(A)^m$ 表示 A 在 m 重笛卡儿积，其中 $A=\prod_{i\in\Gamma} D_i$ 是每个阶段可能的行动组合所组成的集。对于任一 k 和 $(A)^{2k-1}$ 中的任一 $(a^1,\cdots,a^{k-1},c^1,\cdots,c^k)$，数 $u^k(c^k \setminus a^1,\cdots,a^{k-1},c^1,\cdots,c^{k-1})$ 表示在以前阶段所推荐的历史为 $(c^1,\cdots,c^{k-1})$ 和过去行动历史为 $(a^1,\cdots,a^{k-1})$ 的条件下，$c^k=(c_i^k)_{i\in\Gamma}$ 为调解人在阶段 k 推荐给局中人的行动组合的条件概率。对于每个 k，如果推荐是按照相关战略 u 生成的，且每个局中人总是服从推荐，则令 $P^k(a^k\mid u)$ 是在阶段 k 所执行的行动组合为 a^k 的概率。为使其形式上更明确，我们可归纳地定义

$$Q^k(a^1,\cdots,a^k\mid u)=u^k(a^k\mid a^1,\cdots,a^{k-1},a^1,\cdots,a^{k-1})Q^{k-2}(a^1,\cdots,a^{k-1}\mid u)$$

其中 $Q^1(a^1)=u^1(a^1)$，再令

$$P^k(a^k\mid u)=\sum_{(a^1,\cdots,a^{k-1})\in A^{k-1}} Q^k(a^1,\cdots,a^k\mid u)$$

当每个阶段所推荐行动都被公开宣布时，局中人 i 的一个“操纵战略”（manipulative strategy）是任一 $e=(e^k)_{k=1}^{\infty}$，其中对每个 k，e^k 都是从 $(A)^{2k-1}$ 到 A_i 的一个函数。这里 $e^k(a^1,\cdots,a^{k-1},c^1,\cdots,c^k)$ 表示，若 $(a^1,\cdots,a^{k-1})$ 是局中人在阶段 k 之前实际所选择的过去行动历史，且 $(c^1,\cdots,c^k)$ 是直到阶段 k 的所有阶段所推荐的行动历史，则它就是局中人 i 按照战略 e 在阶段 k 所选择的行动。如果所推荐的行动是按照 u 决定的，除 i 之外的每个局中人都服从推荐，但 i 使用了操纵战略 e，则令 $P_i^k(a^k\mid u,e)$ 是局中人在阶段 k 实际所选择的行动组合为 a^k 的概率。为了使其形式上更明确，令

$$Q_i^1(a^1,c^1\mid u,e)=u(c^1)\quad 若\ a^1=(c_{-i}^1,e^1<c^1)$$

$$Q_i^1(a^1,c^1\mid u,e)=0\quad 若\ a^1\neq(c_{-i}^1,e^1(c^1))$$

并对所有 $k>1$，归纳地定义

$$Q_i^k(a^1,\cdots,a^k,c^1,\cdots,c^k\mid u,e)=0，若\ a^k\neq(c_{-i}^k,e^k(a^1,\cdots,a^{k-1},c^1,\cdots,c^k))，$$

$$Q_i^k(a^1,\cdots,a^k,c^1,\cdots,c^k\mid u,e)=u^k(c^k\mid a^1,\cdots,a^{k-1},c^1,\cdots,c^{k-1})$$

$$Q_i^{k-1}(a^1,\cdots,a^{k-1},c^1,\cdots,c^{k-1}\mid u,e)，若\ a^k=(c_{-i}^k,e^k(a^1,\cdots,a^{k-1},c^1,\cdots,c^k))$$

然后令

$$P_i^k(a^k\mid u,e)=\sum_{(a^1,\cdots,a^{k-1},c^1,\cdots,c^k)\in A^{2k-1}} Q_i^k(a^1,\cdots,a^k,c^1,\cdots,c^k\mid u,e)$$

为了描述期望贴现平均支付，令

$$U_i(u,\delta)=(1-\delta)\sum_{k=1}^{\infty}(\delta)^{k-1}\sum_{a\in A}P^k(a\mid u)u_i(a)$$

$$U_i^*(u,e,\delta)=(1-\delta)\sum_{k=1}^{\infty}(\delta)^{k-1}\sum_{a\in A}P_i^k(a\mid u,e)u_i(a)$$

相关战略 u 是标准重复博弈在贴现平均支付准则下的一个“公开相关均衡”（publicly correlated equilibrium）的充分条件是，对于每个局中人 i 和 i 的每个操纵战略

e，都有

$U_i(u,\delta) \ge U_i^*(u,e,\delta)$

对于任一k和$(A)^{2k}$中的任一$(a^1,\cdots,a^k,c^1,\cdots,c^k)$，定义

$u\backslash(a^1,\cdots,a^k,c^1,\cdots,c^k)$使得$u\backslash(a^1,\cdots,a^k,c^1,\cdots,c^k)=(\hat{u}^m)_{m=1}^{\infty}$，当且仅当对于每个$m$和$A^{2m-1}$中的每个$(\bar{a}^1,\cdots,\bar{a}^{m-1},\bar{c}^1,\cdots,\bar{c}^m)$均有

$\hat{u}^m(c^m \mid \bar{a}^1,\cdots,\bar{a}^{m-1},\bar{c}^1,\cdots,\bar{c}^{m-1}) = u^{m+k}(c^m \mid a^1,\cdots,a^k,\bar{a}^1,\cdots,\bar{a}^{m-1},c^1,\cdots,c^k,\bar{c}^1,\cdots,\bar{c}^{m-1})$

因此，如果一个调解人使用相关战略 u，那么$u\backslash(a^1,\cdots,a^k,c^1,\cdots,c^k)$是其在跟随推荐历史$(c^1,\cdots,c^k)$和实际行动历史$(a^1,\cdots,a^k)$之后的子博弈中将使用的相关战略。一个公开相关均衡是子博弈精炼的，当且仅当对于每个k和$(A)^{2k}$中的每个$(a^1,\cdots,a^k,c^1,\cdots,c^k)$，相关战略$u\backslash(a^1,\cdots,a^k,c^1,\cdots,c^k)$仍是一个公开相关均衡。

令$\hat{v}_i$是在仅考虑纯战略时局中人i在单阶段博弈中的最小最大值，即

$\hat{v}_i = \min_{a_{-i}\in\prod_{j\in\Gamma\backslash i}A_j} \max_{a_i\in A_i} u_i(a)$

令F表示相应的单阶段博弈中用相关战略所能达到的期望支付向量集，即$x=(x_i)_{i\in\Gamma}$属于F的充要条件是$\Delta(A)$中存在某个η使得

$x_i = \sum_{a\in A}\eta(a)u_i(a), \quad \forall i\in\Gamma$

于是，我们可以给出如下无名氏定理。

定理 5.5　（Fudenberg 和 Maskin，1986）

对于任一具有有界支付的标准重复博弈，设x是F中任一向量使得集$F\bigcap\{z\in R^N \mid \hat{v}_i \le z_i \le x_i, \forall i\in\Gamma\}$相对于$R^N$有非空的内部，则存在某个数$\bar{\delta}$，使得$0<\bar{\delta}<1$，且对于满足$\bar{\delta}\le\delta<1$的每个$\delta_1$都存在一个相关战略$u$，使得$u$是该重复博弈在贴现平均支付准则下的一个子博弈精炼的公开相关均衡，并且

$U_i(u,\delta) = x_i, \quad \forall i\in\Gamma$

证明：由于支付是有界的，所以存在某个数Ω，使得对于每个i和a，均有$-\Omega < u_i(a) < \Omega$。$F\bigcap\{z\in R^N \mid \hat{v}_i \le z_i \le x_i, \forall i\in\Gamma\}$相对于$R^N$有非空内部的假定意味着，存在$F$中的某个$y$和某个正数$\varepsilon$使得

$\hat{v}_i < y_i - \varepsilon < y_i \le x_i, \quad \forall i\in\Gamma$

并且对Γ中每个j，向量$(y_{-j}, y_j - \varepsilon)$仍在$F$中。

相关战略u的结构最好用$2|N|+1$个行为模式描述：一个初始模式，以及对每个局中人i都有两个模式，一个被我们称之为“对i严厉惩罚”模式，另一个被我们称之为“对i温和惩罚”模式。在初始模式中，调解人在每个回合推荐的行动，都是按照$\Delta(A)$中的一个概率分布使每个局中人j都得到期望支付x_j来决定的。当模式是对i严厉惩罚时，调解人向除i之外的所有局中人推荐其应使用来对付局中人i的最小化最大战略，以达到上述$\hat{v}_i$之定义式中的最小值，而局中人i应使用其最佳反应。当模式是对i温和惩罚时，

调解人在每个阶段推荐的行动是按照$\Delta(A)$中的一个概率分布使局中人i得到期望支付y_{-i}，而使每个其他局中人j都得到期望支付y_j。

调解人停留在初始模式直至某个阶段恰有一个局中人不服从他做出的推荐。在任何一个有且仅有一个局中人不服从调解人推荐的阶段之后，调解人打算做出的推荐是，"在紧接着的k个阶段中的每个阶段都对此局中人严厉惩罚，其后的每个阶段对此局中人温和惩罚"。这些调解计划只是在一个局中人不服从推荐的某个阶段之后才有所改变，在这种情况下，上述的那句话为调解人决定新的调解计划。

在任一阶段，如果局中人i不服从调解，而其余局中人总是服从，那i的贴现平均支付不大于

$$(1-\delta)(\Omega+(\delta)^1\hat{v}_i+\cdots+(\delta)^k\hat{v}_i+(\delta)^{k+1}(y_i-\varepsilon)+(\delta)^{k+2}(y_i-\varepsilon)+\cdots)$$
$$=(1-\delta)\Omega+(\delta-(\delta)^{k+1}P)\hat{v}_i+(\delta)^{k+1}(y_i-\varepsilon)$$

当模式是对i严厉惩罚时，由于在当前回合推荐给i的行动是其对所有其余局中人使用的最小最大化战略的最佳反应，并且不服从只会推迟其计划的期望支付从$\hat{v}_i$增至$y_i-\varepsilon$的时间，故局中人i不可能通过单方面不服从推荐而受重。因为$x_i \geq y_i \geq y_i-\varepsilon$，所以局中人$i$在初始模式和对其他某个局中人严厉惩罚的模式中要比在对他自己严厉惩罚的模式中更好。因而，为了验证 $\underline{u}$ 是一个子博弈精炼的公开相关均衡，只需证实局中人 i 不会在模式是对 i 温和征罚或对某个其他局中人严厉惩罚时想要偏离即可。

当模式是对i温和惩罚时，如果没有局中人不服从，那么对i的未来支付的期望贴现平均支付$y_i-\varepsilon$。当模式是对其他某个局中人严厉惩罚时，i未来支付的期望贴现平均不小于

$$(1-\delta)(-\Omega-(\delta)^1\Omega-\cdots-(\delta)^{k-1}\Omega+(\delta)^k y_i+(\delta)^{k+1}y_i+\cdots)$$
$$=-\Omega(1-(\delta)^k)+(\delta)^k y_i$$

因而，若对每个局中人i，均有

$$y_i-\varepsilon \geq (1-\delta)\Omega+(\delta-(\delta)^{k+1})\hat{v}_i+(\delta)^{k+1}(y_i-\varepsilon)$$

和

$$-\Omega(1-(\delta)^k)+(\delta)^k y_i \geq (1-\delta)\Omega+(\delta-(\delta)^{K+1})\hat{v}_i+(\delta)^{k+1}(y_i-\varepsilon)$$

则 u 是一个子博弈精炼相关均衡。当

$$\frac{y_i-\varepsilon-\hat{v}_i}{\Omega-\hat{v}_i} \geq \frac{1-\delta}{1-(\delta)^{k+1}}$$

时，上述不等式成立。这个不等式的左边是严格为正，而右边在δ收敛于 1 时收敛于$\frac{1}{(k+1)}$。让我们选取 k，使得

$$\frac{y_i-\varepsilon-\hat{v}_i}{\Omega-\hat{v}_i} > \frac{1}{k+1}$$

那么，存在某个$\hat{\delta}$，使得只要$\bar{\delta} \leq \delta < 1$，就有

$y_i-\varepsilon \ge (1-\delta)\Omega+(\delta-(\delta)^{K+1})\hat{v}_i+(\delta)^{k+1}(y_i-\varepsilon)$

又注意到，固定 k 而令 δ 收敛于 1，式

$-\Omega(1-(\delta)^k)+(\delta)^k y_i \ge (1-\delta)\Omega+(\delta-(\delta)^{K+1})\hat{v}_i+(\delta)^{k+1}(y_i-\varepsilon)$ 的左边收敛于 y_i，而右边收敛于 $y_i-\varepsilon$。因而一定存在某个 $\bar{\delta}$，使得只要 $\bar{\delta} \le \delta < 1$，上面两式就同时满足。

证毕！

值得注意的是，由于惩罚行为对惩罚者带来的损害可能大于对被惩罚者带来的损害，所以在局中人 i 迫使某个其他局中人 j 降到其最小最大值的那个阶段，i 可能自己会得到小于最小最大值的支付。于是，在一个局中人偏离 i 后，对他的惩罚可能是不可靠的，除非给予惩罚者某种激励去实施惩罚。在证明过程中所描述的均衡里，对施惩者在严厉惩罚中的努力表现的奖励就是不会成为随后温和惩罚的对象。$F\bigcap\{z\in R^N \mid \hat{v}_i \le z_i \le x_i, \forall i\in\Gamma\}$ 相对于 R^N 有非空内部的假设意味着，不同的局中人的支付之间有足够的独立性，因此可以在对一个局中人进行温和惩罚时，不致于把另外的某个局中人推向其受温和惩罚时的支付水平。在纯战略最小最大值 $\hat{v}_i$ 一般要大于前面所定义的随机最小最大值 v_i。利用调解人可以秘密地把推荐传送给不同的局中人（因而在每个阶段，没有一个局中人知道调解人目前向其他局中人推荐的行动）这一假设，我们就能证明定理 5.5 的一个翻版，其中用 v_i 代表 $\hat{v}_i$ 并删掉"公开"一词。证明过程唯一的变化是，在对局中人 i 严厉惩罚的模式中，调解人向 i 之外的局中人所推荐的行动是他们按照实现前述方程中最小值的相关战略那样随机地选择行动。但是，这个向 i 之外的局中人所推荐的行动对局中人 i 本人是隐蔽的。

5.6.4 厂商为何提供高质量产品：一个重复博弈解说

对于大多数产品来说，消费者在购买时是难一下子判断出产品的质量好坏的。如果厂商提供的产品存在质量问题，往往要待消费者使用之后才会发现。当然，一旦消费者在使用产品后发现产品存在质量问题，往往可以通过诉诸法律的方式指控厂商，如果厂商预期到生产低质量产品可能会招致起诉，他们就会注重产品的质量管理，从而尽量减少产品中的低质量比例。但是，许多消费者即使事后发现所购买的产品存在质量问题，也不一定会起诉，因为起诉存在成本。所以，对厂商的产品质量管理产生约束机制的还不仅限于潜在的起诉可能性，我们在本小节将表明，如果厂商关注长期利益，他们与消费者之间进行的是无限次重复博弈，则存在长期的合作均衡即厂商总是生产高质量产品的均衡。

这里，我们构造一个简单的模型来加以说明。这个模型也是无名氏定理的一个简单应用，不过，与前面的模型不同的是，这里我们将无名氏定理扩展到一部分局中人不固定的重复博弈情形。

假定有一个厂商是长期的固定局中人，他不断重复性地提供产品；但消费者是不固定的，每一阶段有一个消费者只买一次。这样，模型的基本假设是，一个无限次重复博弈的局中人由一个固定的厂商和无限个不同的消费者（也可以一部分消费者是相同的，如同一位消费者在不同的阶段决定买与不买）。厂商在每个阶段决定是生产一单位高质量的产品呢还是生产低质量的产品，而每一个阶段有唯一的一个消费者面临是否购买一单位产品的抉择。假定消费者在购买时不知道自己买的产品的质量，但知道所有之前的消费者购买的产品质量。如果消费者不买，他的支付为 0；如果买到高质量产品，其支付为 1；如果买到的是低质量产品，他的支付就为－1。厂商卖出高质量产品，其支付为 1，但若卖出低质量产品，支付就为 2。这里所说的支付都是指阶段支付。表 5.9 给出了这个博弈的阶段博弈支付矩阵。

表 5.9　产品质量博弈

		厂商	
		高质量	低质量
消费者	买	1, 1	-1, 2
	不买	0, 0	0, 0

如果是一次性博弈，则只有一个纳什均衡（不买，低质量），这对应于走街串巷的小贩们难以取得人们信任的情形。因为游走于街坊之间的小贩们知道即使卖低质量产品坑蒙大家，人们也不一定会在事后抓住他，因为他下一次不会再走同样的路线，这就激励小贩们兜售低质量产品。给定小贩们的这种战略，理性的消费者是不愿购买游贩们的产品的。理性消费者更愿意到大商场或坐摊上买东西，为什么呢？这是由于大商场明天不会不见了，坐摊一般也会长期在同一个地方卖产品，于是，如果买回来的产品存在质量问题，明天还可以去找他们索赔。给定这一战略，厂商若要把生意长期做下去，就不会提供低质量产品，而给定厂商提供高质量产品，消费者就会买。下面，我们来证明这是一个精炼均衡。设厂商执行的战略是：开始生产高质量产品，继续生产高质量产品，除非曾经生产过低质量产品；如果上一次生产了低质量产品，之后永远生产低质量产品。

消费者的战略是，第一阶段的消费者决定买；只要厂商不曾生产过低质量产品，随后阶段的消费者继续买；如果厂商曾经生产过低质量产品，之后的消费者不再买。显然，这些战略实际上是冷酷战略。

给定消费者战略，当厂商生产低质量产品时，得到 2 单位短期支付，但之后每阶段支付为 0；若厂商总是生产高质量产品，每阶段支付为 1，贴现后总支付为 $\frac{1}{(1-\delta)}$，其中 δ 为厂商的贴现因子。当 $\frac{1}{(1-\delta)} \geq 2$ 或 $\delta \geq \frac{1}{2}$ 时，厂商就会总生产高质量产品。所以，

若之前不曾生产低质量产品，厂商的最优选择是继续生产高质量产品。如果之前曾生产低质量产品，给定消费者战略，之后消费者不再买，破罐子破摔地继续生产低质量产品是最优的。这样，当厂商具有足够的长期利益关注时（即δ足够大）给定消费者战略，厂商战略是最优的。反过来，给定厂商战略，假设$\delta \geq \frac{1}{2}$。每个消费者只关心单阶段支付，当之前厂商不曾生产低质量产品时，他预期该阶段厂商将继续生产高质量产品，给定厂商生产高质量产品，消费者的最优选择是买。当之前厂商曾生产低质量产品，他预期厂商在该阶段会生产低质量产品，给定厂商提供低质量产品，他的最优选择是不买。于是，给定厂商战略，消费者战略也是最优的。从而构成一个纳什均衡。下面证明该均衡还是精炼的。首先，在之前厂商不曾生产低质量产品的阶段出发的子博弈上，原有战略组合给出的在子博弈上的限制是完全等同于原有的战略组合，从而是子博弈上的纳什均衡。其次，在之前厂商曾生产低质量产品的阶段出发的子博弈上，原有战略组合给出的在子博弈上的限制是，厂商总生产低质量产品，消费者总选择不买，这显然也构成子博弈上的纳什均衡。所以，原有战略组合是一个子博弈精炼纳什均衡。该博弈的子博弈纳什均衡不只一个，如"消费者永远不买，厂商总生产低质量产品"也是一个子博弈纳什均衡。这个均衡实际上意味着这个厂商并不存在。

Simon（1951）和 Kreps（1986）曾用类似的博弈解释为什么存在"企业"，他们认为企业存在的原因之一正是创造了一个"长期的局中人"，从而使δ足够大，使消费者相信其产品质量，反过来消费者的不断购买又激励企业不断生产高质量产品。

这个模型是 Klein 和 Leffler（1981）提出的模型的一种简化形式，它解释为什么消费者愿意去大商店买东西而不太信赖走街串巷的小贩。

注释：

[1] $\widetilde{A}$既可以是空集（即没有“枝”的图或散点图），也可以为整个 A（称为一个“完全的”图）。

[2] 由于我们在这里要求 A 中元素由 V 中任意两个不同的结点排列而成，故这里排除了 $v_1 = v_2$ 且 $m = 2$ 的情形，因此一个路径至少含有两个结点。

[3] 根据“树”的定义知从 r 到 v_i 有唯一的路径。

[4] 该定义也可等价地写成：如果定义 5.1 中的博弈树 T 只有有限个结点，则称 G 为有限博弈。

[5] 与静态博弈不同的是，在动态博弈情形，混合战略实际上可以有两种方式的定义。一种就是按与战略式表述中的已有定义一样，将随机地在纯战略空间中进行战略选择定义为混合战略；而另一种定义是指局中人在每一个信息集上的行动空间中随机地选择行动，从而在所有信息集上都呈现一种随机性的行动选择，这就可构成一种随机意义上的战略。为了使之区别于第一种定义，一些教科书将这种战略称为“行为战略”（如张维迎，1996），而将第一种定义称为“混合战略”。两种不同的概念是有所不同的，可以证明，行为战略空间比混合战略空间大许多，但两者在博弈均衡的结果上却基本上无差别。Kuhn（1953）证明：在完美回忆博弈中，两种概念是等价的。所以，在许多场合实际上对两种概念不加区别地使用，可统称为“混合战略”。

[6] 子博弈的定义只需要求开始的信息集是单结的，也并不限于完美信息博弈。如果子博弈开始的信息集不是单结的，则局中人不知道他自己位于何种位置，难以形成可分析的博弈。

[7] 杨小凯对张维迎将“Perfect”译成“精炼”表示异议。他认为“精炼”应对应英文间“refine”，而建议将“Perfect”译成“完美”。但笔者认为，若将“Perfect”译成“完美”，似乎与目前博弈论还未将动态博弈纳什均衡真正“完美”地进行精炼的现状不相协调，因为即使是现有的大量精炼工作，还未彻底消除纳什均衡的多重性。“完美”一词似乎带有终极成功的色彩。

[8] 显然，要求在所有子博弈上都构成纳什均衡，说明战略不仅在均衡路径上是最优的，而且在非均衡路径上也是最优的。

子博弈精炼均衡贯彻了这样一种思想：即局中人在博弈的每一个时点上都最优化自己的决策，而无论过去历史怎样，这正是博弈论中常用的“序贯（Sequential）理性”的思想。

[9] 据说美国历史上有位有名的政客幼时是一位流落街头的孤儿，平日以在大街上讨钱为生。人们发现一个不解的现象是如果你让他在一个便士与二个便士之间进行选择，他通常会选择要一个便士而不是更多一些的二个便士。于是经常就有人拿一个便士和两个便士让他选择而观其“傻”取乐，也经常有人为证实这种传言而拿钱去试他。果

然一试就灵，大家终于认为这个孤儿脑子有问题。后来有好心人终于问他道：“难道你不知道两个便士比一个便士更多一些吗？你为什么不要两个便士而只要一个便士呢？”他回答道：“如果我有一次去拿两个便士而不是一个便士，那么，今后就不会有人不断拿钱来试验我了，我也不会讨到那怕是一个便士了。”这个故事的真伪暂且不论，它说明在博弈论中，“装傻”也许也是一种可取得较高效用的策略。小说《红岩》中的华子良为摆脱特务的监视装疯卖傻十多年，《三国演义》中的刘备在势弱时与曹孟德煮酒论英雄时借雷声掩饰自己的“枭雄”之志等都是运用这种策略的成功例子。

[10] 当时的义军是兵分两路的。除了项羽与秦军主力决战的这一路义军外，还有实力较弱但利用项羽吸住秦军主力而避开秦朝大军轻松直取咸阳的另一路义军，即沛公刘邦所率义军。实际上是刘邦先于项羽进入咸阳。按事前的协议，先进咸阳的刘邦应为天下之王，但刘邦实际上是借项羽吸引住秦军主力而占了便宜顺利入咸阳的，并且刘邦当时在军力上远非项羽的对手。所以，随后并未遵守协议，项羽成为了天下之王，但后来多年演变出的楚汉之争最终使善于用人的刘邦胜出，刘邦治下的汉军大将韩信（原本是项羽帐下无名之将，因未能任用转投刘邦）在垓下一战彻底击溃项羽（项羽本人也逃到乌江边自刎）。

[11] 毛泽东在 1949 年人民解放军追击溃败中的蒋介石大军时曾在其诗词中反其意而用之，言“宜将剩勇追穷寇，不可沽名学霸王”。这里霸王指楚霸王项羽。项羽正是未在几次自己占绝对优势的场合下杀掉刘邦而养虎为患，最后自己却败在刘邦手下。

[12] 我们将在后面的章节中指出，过剩生产能力还是垄断性企业向潜在竞争者显示自己有竞争能力的信号。因为一旦进入者进入，原有的在位者将通过降低产品价格打击进入者，而拥有较大规模的企业将有较大的规模经济优势，因而可以通过降价增产来打击对手，因为这种企业在产量增加时平均成本会下降，降价不会导致亏损。过剩生产能力或较大的生产规模是向进入者显示自己的“类型”，即自己是强有力的竞争者，否则威胁是不可置信的。

[13] 该博弈还有其它精炼均衡，如每阶段都选“坦白”仍是一个精炼均衡。

[14] 若有 n 个寡头企业，则默契合作均衡要求 $\delta \geq \left[1+\frac{4n}{(n+1)^2}\right]^{-1}$。当 $n \to \infty$，$\delta \to 1$，即企业愈多，默契合作就愈困难（对 δ 要求愈大，要求局中人愈有耐心）。这一结论与现实观察相吻合。它也可解释为什么小团体的合作靠非正式的规则就可维持，而大团体就必须依赖于正式的规则和合约（《三国演义》中刘、关、张桃园结义，三弟兄多年来生死同命，情深谊长；但人数较多的水泊梁山上的 108 将之中却各有小圈子，更有甚者的是瓦岗寨中人数较多的瓦岗义军诸将却同床异梦，互相残杀；另外，在现实社会中的小圈子，铁哥们，都是小集体易合作的例子。）。若回到囚徒困境，则为：干坏事的合伙人愈多，事情就愈易败露。

[15] 与囚徒困境一样，该博弈也有许多精炼均衡，如“总是选古诺均衡产量”也是一个精炼均衡。可证明：任何介于古诺均衡产量和垄断均衡产量间的产量（即 $q \in \left[\frac{q^M}{n}, q_i^c\right]$都是某个精炼均衡的一个特定结果。这是下面将要介绍的无名氏定理的直接推论。其含意为：在现实中，完全对立不合作与完全合作两个极端都是少见的，更多的是介于两个极端之间的“若即若离”状态（与δ有关）。

第六章　不完全信息静态博弈

博弈论在 1970 年代之后逐渐进入主流经济学体系，主要是由于它在不完全信息条件下的经济分析中表现出一种特别的魅力。不完全信息指经济活动中一部分经济主体的某些特征对于其他主体来说是不清楚的，如在拍卖商品或工程招投标中，参加拍卖的潜在买主愿意为拍卖品支付的最高价格或参加工程招投标的投标者愿意为工程开出的最低价格只能是各个潜在买主或投标者心中的秘密，其他人是不清楚的，即使潜在买主或投标者告诉其他人他们愿支付的最高价格或最低价格，其他人也不会相信他们说的是真的。潜在买主或投标者也知道其他人并不清楚他们愿开出的最高价格或最低价格，因而也不会直接说出真实的价格底线。信息不完全又称为信息不对称，即其他局中人没有特定局中人清楚特定局中人自身的特征。不完全信息静态博弈就是假定某些局中人具有其他局中人不清楚的某些特征的静态博弈。当然，对于局中人本身来说，他自身的这些不为人所知的特征对于他自己来说是清楚的，因而称这些特征为局中人自己拥有的“私人信息”（private information）。在博弈论中，习惯地将局中人的“私人信息”集中表现为局中人的支付函数特征，也就是说，局中人的私人特征将完全通过其支付函数特征表征出来，而不完全信息就表现为一些局中人不清楚另一局中人的支付函数，当然，每个局中人是完全清楚自己的支付函数的。

6.1 不完全信息博弈：基本概念

在假定局中人拥有私人信息的情况下，其他局中人对特定局中人的支付函数类型是不清楚的。如果一些局中人不知道另一些局中人的支付函数，或支付函数不是共同知识，局中人就不知道他在与谁博弈，因而在 1967 年以前，博弈论专家认为此时博弈的结构特征是不确定的，无法进行分析。Harsanyi（1967、1968）提出了一种处理不完全信息博弈的方法，即引入一个虚拟的局中人——“自然 N”。N 首先行动，它决定每个局中人的特征。每个局中人知道自己的特征，但不知道别的局中人特征。这种方法将不完全信息静态博弈变成一个两阶段动态博弈，第一个阶段是自然 N 的行动选择，第二阶段是除 N 外的局中人的静态博弈。这种转换被称为“Harsanyi 转换”，它将不完全信息博弈转换为完全但不完美信息博弈。

称局中人拥有的私人信息为他的“类型”，许多情况下局中人类型由其支付函数完全决定，故常将支付函数等同于类型。用 θ_i 表示局中人 i 的一个特定类型，H_i 表示局中人 i 所有可能类型的集合，即 $\theta_i \in H_i$，称 H_i 为局中人 i 的类型空间，$i = 1, \cdots, n$。

在不完全信息静态博弈中，由于局中人的类型存在多种可能，因而与局中人相关的各种概念都随其类型的不同而不同，其中之一就是局中人的行动空间 A_i 将随类型 θ_i

而变化，即 $A_i = A_i(\theta_i)$，如企业能够选择的产量范围依赖于其成本函数。由于支付函数刻画了类型，或支付函数也是类型依存的，如同样产量不同成本函数的企业的利润就不同。我们可将其记为[1]

$$u_i = u_i(a_1, \cdots, a_i, \cdots, a_n, \theta_i) \tag{6.1}$$

$i = 1, \cdots, n$

在其他局中人已选定行动 a_j， $j = 1, \cdots, n$， $j \neq i$ 时，局中人选行动 $a_i \in A_i(\theta_i)$ 获得的支付由式（6.1）给出。显然给定 a_j 时最大化 u_i 的 a_i 与 θ_i 有关，即 $a_i^* = a_i^*(\theta_i)$，其中 a_i^* 是给定 a_j 时最大化 u_i 的 a_i。这里，我们用“类型依存”来描述包括最优战略在内的概念与类型的对应关系。可以预料到，当局中人的类型给定时，其最优战略 a_i^* 的范围也就是给定了的，当其是唯一存在的时，有 $a_i^* = a_i^*(\theta_i)$。在前面介绍的行动空间 A_i，其类型依存就是 $A_i(\theta_i)$。

对于局中人 i 来说，他不知道其他局中人的类型。当他选择任一行动 $a_i \in A_i(\theta_i)$ 时，对于其他局中人的任一可能的类型组合 $\theta_1, \cdots, \theta_{1-1}, \theta_{1+1}, \cdots, \theta_n$，如果给定其他局中人类型与其最优战略 a_j 的一个对应组合 $a_j^*(\theta_j)$， $j \neq i$， $j = 1, \cdots, n$，其支付为

$$u_i(a_1^*(\theta_1) \cdots a_{i-1}^*(\theta_{i-1}), a_i, a_{i+1}^*(\theta_{i+1}) \cdots a_n^*(\theta_n), \theta_i)$$

假定局中人 i 认为其他局中人的类型组合恰好为 θ_{-i} 的概率为 $P_i(\theta_{-i}|\theta_i)$。这一概念与局中人 i 的类型 θ_i 有关，还与博弈开始之时局中人 i 对 $\theta_1 \cdots \theta_i \cdots \theta_n$ 的概率分布知识有关，记这一事前的分布密度（局中人在博弈之前掌握的知识使其对 $\theta_1 \cdots \theta_i \cdots \theta_n$ 的取值可能性判断）为 $P_i(\theta_1, \cdots \theta_i, \cdots, \theta_n)$，称为“先验概率”。显然， $P_i(\theta_{-i}|\theta_i)$ 就是一种条件概率的概念，即当局中人 i 的类型为 θ_i 时，他认为 θ_{-i} 取值 θ_i 的概率为

$$\begin{aligned} P_i(\theta_{-i}|\theta_i) &= \frac{P(\theta_i, \theta_{-i})}{P(\theta_i)} \\ &= \frac{P(\theta_i, \theta_{-i})}{\sum_{\theta_{-i} \in H_{-i}} P(\theta_{-i}, \theta_i)} \end{aligned} \tag{6.2}$$

式（6.2）是概率论中著名的贝叶斯公式（Bayes equition）。

此时，因为 i 不知道 θ_{-i}，与第 3 章中的混合博弈支付函数构造相类似，我们这里用 von·Neumann——Morganstern 效用函数刻画这种不确定下的支付函数，即 i 的期望支付为[2]

$$\sum_{\theta_{-i}} P_i(\theta_{-i}|\theta_i) u(a_{-i}^*(\theta_{-i}), a_i, \theta_i) \tag{6.3}$$

显然，最大化式（6.3）的 a_i 就是 i 的最优战略 a_i^*，它与 θ_i 有关，故 $a_i^* = a_i^*(\theta_i)$。当存在一组“类型依存”的最优战略 $a^* = (a_1^*, \cdots, a_i^*, \cdots, a_n^*)$，满足

$$a_i^* \in \arg\max_{a_i \in A_i(\theta_i)} \sum_{\theta_{-i}} P_i\{\theta_{-i} \mid \theta_i) u(a_{-i}^*(\theta_{-i}), a_i, \theta_i) \tag{6.4}$$

$i = 1, \cdots, n$ 则称 a^* 是一个（纯战略）纳什均衡，也称为贝叶斯纳什均衡（Bayes Nash

equilibrium）。显然 a^* 是类型依存的，即 $a_i^* = a_i^*(\theta)$ 。为了减少复杂性，假定博弈开始之前各个局中人掌握的关于 $(\theta_1,\cdots,\theta_n)$ 的分布密度知识是相同的，于是有：

Harsanyi 公理：假定概率分布密度 $P(\theta_1,\cdots,\theta_n)$ 是所有局中人的共同知识。这一公理表明所有局中人有关自然行动的信念（belief）是相同的。

贝叶斯纳什均衡（也简称贝叶斯均衡）是完全信息静态博弈纳什均衡概念在不完全信息静态博弈上的扩展。

有时也称不完全信息静态博弈为静态贝叶斯博弈或贝叶斯静态博弈。

在上面，我们实际上已给出了静态贝叶斯博弈的战略式表述，以下给出正式的战略式表述和一些例子。

6.2 静态贝叶斯博弈的战略式表述：贝叶斯纳什均衡

定义 6.1（n 人静态贝叶斯博弈的战略式表述）局中人类型空间 $H_1,\cdots,H_n$ ；条件概率 $P_1,\cdots,P_n$ ；战略空间 $A_1(\theta_1),\cdots,A_n(\theta_n)$ ；支付函数 $u_1(a_1,\cdots,a_n,\theta_1),\cdots,u_n(a_1,\cdots,a_n,\theta_n)$ 。i 知道 θ_i 。用 $G=\{A_1,\cdots,A_n,\theta_1,\cdots,\theta_n,P_1,\cdots,P_n,u_1,\cdots,u_n\}$ 表示该博弈。

博弈顺序为：

①自然 N 选 $\theta=(\theta_1,\cdots,\theta_n)$ ， $\theta_i \in H_i$ ， 局中人 i 观察到 θ_i ，但局中人 $j \neq i$ 仅知道 $P_j(\theta_{-j} \mid \theta_j)$ ，不能观察到 θ_i 。

② n 个局中人同时选行动（战略） $a=(a_1,\cdots,a_n)$ ， $a_i \in A_i(\theta_i)$ 。

③ i 得到支付 $u_i(a_1,\cdots,a_n,\theta_i)$ ， $i=1,\cdots,n$ 。

注意，该定义不排除局中人 j 可能拥有关于局中人 i 类型的某种信息，而当所有局中人类型空间只有一个元素时，不完全信息静态博弈就退化为完全信息静态博弈。

同时，这里假定了 $A_i(\theta_i)$ 和 $u_i(a_i,a_{-i},\theta_i)$ 本身是共同知识，即尽管其他局中人不知道 i 的类型 θ_i ，但他们知道 i 的战略空间和支付函数是如何依赖于他的类型的，即当他们知道 θ_i 时，就必然知道 $A_i(\cdot)$ 和 $u_i(\cdot)$ 。

n 人不完全信息静态博弈 $G=\{A_1,\cdots,A_n,\theta_1,\cdots,\theta_n,P_1,\cdots,P_n,u_1,\cdots,u_n\}$ 的纯战略贝叶斯纳什均衡是一个类型依存战略组合 $\{a_i^*(\theta_i)\}_{i=1}^n$ ，满足[3]：

$$a_i^*(\theta_i) \in \arg\max_{a_i \in A_i(\theta_i)} \sum P_i\{\theta_{-i} \mid \theta_i\} u_i(a_{-i}^*(\theta_{-i}),a_i,\theta_i)$$

$$i=1,\cdots,n$$

可以定义混合战略贝叶斯纳什均衡，并且模仿第 4 章中的方法不难证明有限博弈的不完全信息静态博弈至少存在一个贝叶斯纳什均衡，这一工作可留给读者自己去完成。

需要指出的是：纯战略只是一个行动 a_i 如何依类型 θ_i 而变的规则，不是指一个具

体的结果。

6.3 某些静态贝叶斯博弈的例子

例子 6.1　市场进入博弈

一个完全垄断企业 B 正在垄断一个行业市场，另一个潜在的试图进入该行业的企业 A，称 A 为进入者，B 为在位者。A 不知道 B 的成本特征，设 B 有两种可能的成本，即高成本和低成本。两种成本情况下的博弈矩阵如表 6.1。

表 6.1　市场进入博弈

		B			
		高成本		低成本	
		默认	斗争	默认	斗争
A	进入	40,50	-10,0	30,80	-10,100
	不进入	0,300	0,300		0, 400

假定 B 知道进入者 A 的成本为高成本，且与 B 为高成本时的成本相同。假若信息是完全的，则当 B 为高成本时，唯一的精炼纳什均衡为（进入，默认），另一纳什均衡（不进入，斗争）是含有不可置信的威胁。当 B 为低成本时，唯一的纳什均衡为（不进入，斗争），即若 A 进入行业，具有低成本优势的 B 将通过降低价格将 A 逐出市场。由于存在行业进入成本，所以 A 被逐出市场后将有净的 10 单位进入成本的损失。

当 A 不知道 B 的成本情况时，他的选择将依赖于他对 B 的成本类型的主观概率或先验概率密度。

设 A 对 B 是高成本的先验概率判断为 P，则 A 认为 B 为低成本的概率为 $1-P$。

如果 A 进入，其期望支付为

$$P(40)+(1-P)(-10)$$

如果 1 不进入，其期望支付为 0。

当且仅当 $P(40)+(1-P)(-10)\geq 0$ 或 $P\geq\frac{1}{5}$ 时，A 选择进入；反之，当 $P<\frac{1}{5}$ 时，A 不进入。

于是，贝叶斯均衡为：

（进入，默认），高成本，$P\geq\frac{1}{5}$；

（进入，斗争），低成本，$P\geq\frac{1}{5}$；

（不进入，*），$P<\dfrac{1}{5}$

其中*表示可以是斗争，也可以是默认。

例 6.2 成本信息不对称的古诺博弈

例 3.10 给出的古诺博弈中，每个厂商的成本函数是共同知识。这里，我们假设每个厂商的成本函数是私人信息，具体规定如下：两个企业生产相同产品在同一市场上进行竞争性销售，市场需求函数为$Q=a-P$，$a>0$，P为产品价格，Q 为市场需求量。假设a充分大时总有$a-P\geq 0$，企业i的成本函数为$C_i=b_iq_i$，其中C_i为企业i的总成本，q_i为其产量，b_i为其平均成本，b_i为常数且$b_i>0$，故b_i也是边际成本。b_i是企业i的私人信息，企业j不知道b_i但认为b_i在$[d,e]$上呈均匀分布，$d>0$，$e>0$，$d\leq e$。且进一步假定b_i在$[d,e]$呈均匀分布是共同知识，$i\neq j$，$i=j=1,2$。

企业i的支付函数是其利润函数π_i

$$\begin{aligned}\pi_i&=Pq_i-c_i\\&=(a-Q)q_i-b_iq_i\end{aligned}$$

因 $Q=q_1+q_2$

故 $\pi_i=(a-q_1-q_2)q_i-b_iq_i$

设静态贝叶斯均衡为$\{q_i^*\}_{i=1,2}$，则由均衡战略的类型依存性有

$$q_i^*=q_i^*(b_i),\ \ i=1,2$$

于是

$$\begin{aligned}\pi_i&=(a-q_1^*(b_1)-q_2^*(b_2))q_i^*(b_i)-b_iq_i^*(b_i)\\&=\pi_i(b_j)\end{aligned}$$

i的期望支付为

$$u_i=\int_{H_j}P_i(b_j\,|\,b_i)\pi_i(b_j)db_j$$

显然$P(b_j\,|\,b_i)=P(b_j)$，由概率分布密度$P(b_j)$的归一化条件

$$\int_{H_j}P(b_j)db_j=1$$

及b_j在$[d,e]$上呈均匀分布假设，有

$$P(b_j)\int_{H_j}db_j=1$$

或$[e-d]P(b_j)=1$

即$P(b_j)=\dfrac{1}{e-d}$

于是，$u_i=\dfrac{1}{e-d}\left\{\left[(a-q_i)(e-d)-\int_{H_j}q_j(b_j)db_j\right]q_i-b_iq_i(e-d)\right\}$

一阶条件：

$$\frac{\partial u_i}{\partial q_i}=\frac{1}{(e-d)}\left[-(e-d)q_i+(a-q_i)(e-d)-\int_{H_j}q_j(b_j)db_j-b_i(e-d)\right]=0$$

即 $q_i=\dfrac{(a-b_i)(e-d)-\int q_j}{2(e-d)}$ （6.5）

同样由对称性有

$$q_j=\frac{(a-b_j)(e-d)-\int q_i}{2(e-d)} \quad (6.6)$$

在上式两端对 b_j 进行积分

$$\int q_j=\frac{a(e-d)-\frac{e^2-d^2}{2}-\int q_i}{2} \quad (6.7)$$

在式（6.5）两端对 b_i 积分

$$\int q_i=\frac{a(e-d)-\frac{e^2-d^2}{2}-\int q_j}{2} \quad (6.8)$$

将式（6.7）代入式（6.8）的右端，得

$$\int q_i=\frac{a-\frac{e+d}{2}}{3}(e-d) \quad (6.9)$$

由对称性有 $\int q_j=\int q_i=\dfrac{a-\frac{e+d}{2}}{3}(e-d)$

代入式（6.5）得

$$q_i^*=\frac{(a-b_i)(e-d)-\frac{a-\frac{e+d}{2}}{3}(e-d)}{2(e-d)}$$

$$=\frac{2a-3b_i+\frac{e+d}{2}}{6}$$

同理有 $q_j^*=\dfrac{2a-3b_j+\frac{e+d}{2}}{6}$

于是得静态贝叶斯均衡为 $\left(\dfrac{2a-3b_1+\frac{e+d}{2}}{6},\dfrac{2a-3b_2+\frac{e+d}{2}}{6}\right)$。

当 a 充分大时，q_i^* 和 q_j^* 均为非负数。

当 $b_1\ge b_2$ 时，$q_1^*<q_2^*$；

均衡利润 $\pi_1^*=(P-b_1)q_1^*<(P-b_2)q_2^*=\pi_2^*$，即成本较高的一方利润较低，产量较低。

当$e=d$时，博弈退化成完全信息静态博弈的场合。为了与例 3.26 相比较，进一步设$d=e=c$，$b_1=b_2=c$，则

$$q_1^*=q_2^*=\frac{1}{3}(a-c)$$

这正好回到例 3.26 的结果。

若假设$d<c<e$，$c=\frac{d+e}{2}$，$b_1=b_2=c$，则

$q_1^*=q_2^*=\frac{1}{3}(a-c)$，这与完全信息博弈均衡相同。

若假设$c<\frac{d+e}{2}$，$b_1=b_2=c$，则

$q_1^*=q_2^*>\frac{1}{3}(a-c)$，此时每个厂商都误以为对方的成本较自己高的可能性大一些，从而过于自信地扩大产量。

相反，若假设$c>\frac{d+e}{2}$，$b_1=b_2=c$，则

$q_1^*=q_2^*<\frac{1}{3}(a-c)$，此时每个厂商都误以为对方的成本较自己低的可能性大一些，从而过于谨慎地计划自己的产量。

例 6.3　不完全信息市场交易模型

本例介绍一个模型，它描述一个潜在的商品卖者与一个潜在买者在信息不完全情况下的交易可能性及交易价格的决定。这个模型表达的现实交易内容是广泛的，如可以应用于企业招聘工人的场合（如 Hall 与 Lazear（1984））。

假设买方对商品的估价为v_b，卖方的估价为v_s，双方的估价都是私人信息，并且服从$[0,1]$区间的均匀分布。博弈规则是：卖方确定一个卖价P_s，买方同时给出一个买价P_b。当$P_b\ge P_s$，则交易达成且价格为$P=\frac{P_b+P_s}{2}$；当$P_b<P_s$，则不发生交易。当交易达成时，买方的支付为v_b-P，否则支付为 0；交易达成时卖方支付为$P-v_s$，否则为 0。

设静态贝叶斯均衡为$\{P_b(v_b),P_s(v_s)\}$，给定$v_b, P_b=P_b(v_b)$是最大化如下期望支付的解：

$$\max_{P_b}\int_{P_b\ge P_s(v_s)}P(v_s\mid v_b)[v_b-\frac{P_b+P_s(v_s)}{2}]dv_s$$

该条件可表为

$$\max_{P_b}\left\{v_b\int_{P_b\ge P_s(v_s)}P(v_s\mid v_b)dv_s-\frac{P_b\int_{P_b\ge P_s(v_s)}P(v_s\mid v_b)dv_s+P_s\int_{P_b\ge P_s(v_s)}P(v_s\mid v_b)P_s(v_s)dv_s}{2}\right\}$$

$$= \max_{P_b}\left\{ v_b\, prob(P_b \ge P_s(v_s)) - \frac{\left(P_b\, prob(P_b \ge P_s(v_s)) + \frac{\int_{P_b \ge P_s(v_s)} P(v_s \mid v_b) P_s(v_s) dv_s}{prob(P_b \ge P_s(v_s))} prob(P_b \ge P_s(v_s)\right)}{2} \right\}$$

$$= \left\{ \max_{P_b}\left[v_b - \frac{P_b + E[P_s(v_s) \mid P_b \ge P_s(v_s)]}{2} \right] prob(P_b \ge P_s(v_s) \right\} \qquad (6.10)$$

其中 $E[P_s(v_s) \mid P_b \ge P_s(v_s)]$ 是在卖方价格小于买方价格条件下，卖方价格的期望值。类似地，给定 v_s， $P_s = P_s(v_s)$ 是最大化如下期望支付的解：

$$\max_{P_s}\left[\frac{P_s + E[P_b(v_b) \mid P_b(v_b) \ge P_s]}{2} - v_s \right] prob[P_b(v_b) \ge P_s] \qquad (6.11)$$

其中 $E[P_b(v_b) \mid P_b(v_b) \ge P_s]$ 为在买方价格大于卖方价格 P_s 的条件下，买方价格的期望值。

该模型不仅存在贝叶斯均衡，而且均衡是十分多的。下面给出一个均衡，称为“单一价格均衡”。此均衡在交易达成时，交易价格就只是单一的一种价格。在这个均衡中， $P_b(v_b)$ 和 $P_s(v_s)$ 的规定如下：

令 $x \in [0,1]$

$$P_b(v_b) = \begin{cases} x & v_b \ge x \\ 0 & v_b < x \end{cases}$$

$$P_s(v_s) = \begin{cases} x & v_s \le x \\ 1 & v_s > x \end{cases}$$

我们来证明如此构造的 $P_b(v_b)$ 和 $P_s(v_s)$ 构成一个贝叶斯均衡。

给定买方战略，若 $v_s \le x$，当 $P_s > x$ 时，不会成交，因为买方价格 P_b 不会超过 x，而卖方报出 $P_s = x$ 时，在 $v_b \ge x$ 时可以成交且以 $P = x$ 售出，此时卖方支付为 $x - v_s \ge 0$，当 $P_s < x$ 时，成交时必有 $v_b \ge x$ 且 $P_b = x$，成交价为 $\frac{P_s + x}{2} < x$，此时卖方支付没有报 $P_s = x$ 的成交价 x 高，所以，此时 $P_s = x$ 是最优的。若 $v_s > x$，当 $P_s > x$ 时，不会成交，而 $P_s \le x$ 时即使成交卖方支付也不会为正。所以，此时卖方的最好选择是不成交，可以令 $P_s = 1$，此时不会成交。于是，给定 $P_b(v_b)$ 如上规定， $P_s(v_s)$ 是卖方的最优战略。

给定 $P_s(v_s)$，下面证明如上规定的 $P_b(v_b)$ 是买方的最优战略。

若 $v_b \ge x$，买方报出 $P_b > x$ 时在 $v_s \le x$ 时会以价格 $\frac{P_b + x}{2} > x$ 成交，但不如报出 $P_b = x$ 好，因此时将以 x 价格成交，当 $v_s > x$ 时，若 $P_b = 1$ 会以价格 1 成交，但买方支付此时不会为正，不如报 $P_b = x$ 不成交为好。所以，当 $v_b \ge x$ 时，买方报 $P_b = x$ 比报 $P_b > x$ 为好。若 $P_b < x$，不会成交。当 $P_b = x$ 时，在 $v_s \le x$ 时成交，买方支付为 $v_b - x \ge 0$。

所以，当 $v_b \geq x$ 时，$P_b = x$ 是最优的。当 $v_b < x$ 时，买方只有报 $P_b \geq x$ 才会成交，但此时买方支付 $v_b - P_b < 0$，所以不成交为好。令 $P_b = 0$ 就可达到不成交的目的。所以，给定卖方战略，买方战略是最优的。

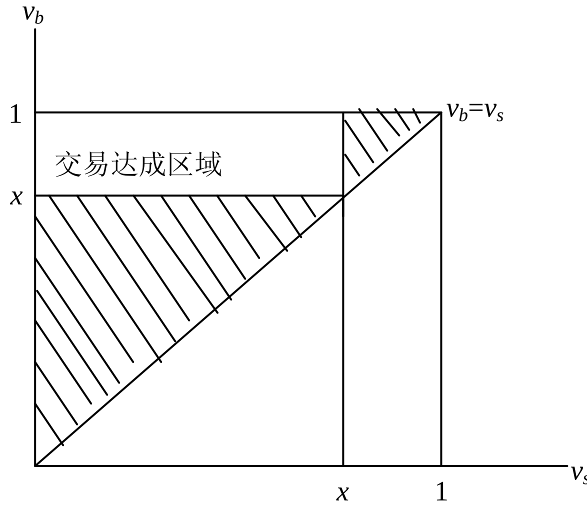


图 6.1　单一价格均衡的交易达成区域

在图 6.1 中，满足 $P_b \geq P_s$ 的区域即交易达成区域仅是有效率交易区域 $v_b \geq v_s$ 的一部分，但其中阴影部分尽管是有效率的交易区域但却不能达成交易。

下面，我们尝试找出另外的一些均衡，譬如最优战略都是线性函数的情形。之所以专门考察线性战略的情形，是因为线性均衡有着如下将要表述的十分有趣的效率特征。

设卖方的战略为 $P_s(v_s) = a_s + c_s v_s$，则 P_s 服从区间 $[a_s, a_s + c_s]$ 上的均匀分布，此时当 $a_s \leq P_b \leq a_s + c_s$, $prob[P_b \geq P_s(v_s)] = prob(P_b \geq a_s + c_s v_s)$

$$= prob(v_s \leq \frac{P_b - a_s}{c_s})$$

$$= \frac{P_b - a_s}{c_s}$$

当 $P_b < a_s$，$prob(P_b \geq P_s(v_s)) = 0$

当 $P_b \geq a_s + c_s$，$prob(P_b \geq P_s(v_s)) = 1$

当 $P_b < a_s$ 时，$prob(P_b \geq P_s(v_s)) = 0$，期望支付也为 0。

令函数 $f(P_b, V_b) = \left[V_b - \frac{1}{2}\left[P_b + \frac{a_s + P_b}{2}\right]\right]\frac{P_b - a_s}{c_s}$

则 $\frac{\partial f(P_b, V_b)}{\partial P_b} = \frac{-6P_b + 2a_s + 4V_b}{4c_s}$

令 $\frac{\partial f}{\partial P_b}=0$，则 $P_b=\frac{2}{3}V_b+\frac{1}{3}a_s$

当 $P_b<\frac{2}{3}V_b+\frac{1}{3}a_s$，则 $\frac{\partial f}{\partial P_b}>0$， f 是 P_b 的严格增函数

当 $P_b>\frac{2}{3}V_b+\frac{1}{3}a_s$， $\frac{\partial f}{\partial P_b}>0$， f 是 P_b 的严格减函数

故 $f(P_b,V_b)$ 在 $\frac{2}{3}V_b+\frac{1}{3}a_s$ 处达到最大值。

1）若 $\frac{2}{3}V_b+\frac{1}{3}a_s\le a_s\Leftrightarrow V_b\le a_s$

则当 $P_b<a_s$ 时， $\Pr ob\left[P_b\ge P_s(v_s)\right]=0$

期望支付也为 0。

当 $P_b>a_s+c_s$ 时，期望支付为

$V_b-\frac{1}{2}\left[P_b+a_s+c_s/2\right]$

它是 P_b 的减函数，在 $P_b=a_s+c_s$ 处为

$V_b-a_s-3c_s/4$

当 $a_s\le P_b\le a_s+c_s$

期望支付 $=f(P_b,V_b)$

在 $P_b=a_s$ 处， $f=\left[V_b-\frac{1}{2}\left[a_s+\frac{a_s+a_s}{2}\right]\right]\times 0=0$

在 $P_b\ge a_s+c_s$ 处，较大的一个支付为

$V_b-a_s-3c_s/4\le -3c_s/4<0$（设 $c_s>0$）

故所有 $P_b\in\left[0,a_s\right]$ 都是最优报价，期望支付为 0，故最优报价可取 $P_b=\frac{2}{3}V_b+\frac{1}{3}a_s$。

2）若 $a_s\le\frac{2}{3}V_b+\frac{1}{3}a_s\le a_s+c_s$

$$\Leftrightarrow \quad a_s\le V_b\le a_s+\frac{3}{2}c_s$$

此时 f 在 $\frac{2}{3}V_b+\frac{1}{3}a_s$ 处达到最大值 $\frac{(V_b-a_s)^2}{3c_s}\ge 0$，而当 $P_b>a_s+c_s$ 在 P_b 趋于 a_s+c_s 时的支付为

$V_b-a_s-3c_s/4$

因为有 $\frac{(V_b-a_s)^2}{3c_s}\ge V_b-a_s-\frac{3c_s}{4}$

这是由于上式等价于 $(V_b-a_s)^2-3c_s(V_b-a_s)+\frac{ac_s^2}{4}\ge 0$

或$\left[(V_b - a_s) - \frac{3}{2}c_s\right]^2 \geq 0$

这是显然成立的。

故最优报价为

$$P_b = \frac{2}{3}V_b + \frac{1}{3}a_s$$

3）若$\frac{2}{3}V_b + \frac{1}{3}a_s > a_s + c_s$

$$\Leftrightarrow \ V_b > a_s + \frac{3}{2}c_s$$

这时若$P_b > a_s + c_s$，则P_b趋于$a_s + c_s$时，期望支付为

$V_b - a_s - 3c_s/4$

而当$P_b = a_s + c_s$时，期望支付为

$$f(P_b, V_b) = \left[V_b - \frac{1}{2}\left[a_s + c_s + \frac{a_s + a_s + c_s}{2}\right]\right]\cdot 1$$

$$= V_b - \frac{1}{2} - \frac{4a_s + 3c_s}{2}$$

$$= V_b - \frac{4a_s + 3c_s}{4}$$

$$= V_b - a_s - 3c_s \ / \ 4 > \frac{3c_s}{4} > 0$$

故此时最优报价为$P_b = a_s + c_s$。

于是在给定$P_s(V_s) = a_s + c_s V_s$下，买方的类型依存战略为

$$P_b(V_b) = \begin{cases} \frac{2}{3}V_b + \frac{1}{3}a_s & 0 \leq V_b < a_s + \frac{3}{2}c_s, \quad \left(a_s + \frac{3}{2}c_s < 1\right) \\ a_s + c_s & 1 \geq V_b \geq a_s + \frac{3}{2}c_s \end{cases}$$

为了获得线性解，我们假设$a_s + \frac{3}{2}c_s \geq 1$，则

$$P_b(V_b) = \frac{2}{3}V_b + \frac{1}{3}a_s \qquad (6.12)$$

以下再看给定买方的上述战略，卖方的最优战略是什么。

给定V_s， $P_s = P_s(V_s)$是最大化如下期望支付的解：

$$\max_{P_s}\left[\frac{P_s + E[P_b(V_b)/P_b(V_b) \geq P_s}{2} - V_s\right]\Pr ob\left[P_b(V_b) \geq P_s\right]$$

因为$P_b(V_b)$是$\left[\frac{1}{3}a_s, \frac{1}{3}a_s + \frac{2}{3}\right]$上的均匀分布。

当$P_s \prec \frac{1}{3}a_s$时， $\Pr o\left[P_b(V_b) \geq P_s\right] = 1$

$$E\left[P_b(V_b)\middle|P_b(V_b)\ge P_s\right]=\frac{1}{2}\left[\frac{1}{3}a_s+\left(\frac{1}{3}a_s+\frac{2}{3}\right)\right]=\frac{1}{3}a_s+\frac{1}{3}$$

期望支付为$\left[\frac{P_s+\left(\frac{1}{3}a_s+\frac{1}{3}\right)}{2}-V_s\right]\cdot 1=\frac{P_s+\frac{1}{3}(a_s+1)}{2}-V_s$

它是P_s的增函数，在$P_s=\frac{1}{3}a_s$，期望支付为$\frac{1}{3}a_s+\frac{1}{6}-V_s$，当$P_s>\frac{1}{3}a_s+\frac{2}{3}$时，$\Pr o\left[P_b(V_b)\ge P_s\right]=0$，期望支付为零。当$\frac{1}{3}a_s\le P_s\le\frac{1}{3}a_s+2/3$时，

$$\Pr o\left[P_b(V_b)\ge P_s\right]=\frac{\frac{1}{3}a_s+\frac{2}{3}-P_s}{2/3}$$

$$E\left[P_b(V_b)\middle|P_b(V_b)\ge P_s\right]=\frac{P_s+\frac{1}{3}a_s+2/3}{2}$$

期望支付为$\left[\frac{P_s+\frac{P_s+\frac{1}{3}a_s+\frac{2}{3}}{2}}{2}-V_s\right]\frac{\frac{1}{3}a_s+\frac{2}{3}-P_s}{2/3}$

$$=\frac{\left[3P_s+\frac{1}{3}a_s+\frac{2}{3}-4V_s\right]}{8/3}\left(\frac{1}{3}a_s+\frac{2}{3}-P_s\right)$$

$$=\frac{\left(3P_s+\frac{1}{3}a_s+2/3-4V_s\right)}{8}(a_s+2-3P_s)$$

$$=g(P_s)$$

$$\frac{\partial g}{\partial P_s}=\frac{3}{8}(a_s+2-3P_s)+\frac{\left(3P_s+\frac{1}{3}a_s+2/3-4V_s\right)}{8}(-3)$$

$$=2a_s+4+12V_s-18P_s$$

令$\frac{\partial g}{\partial P_s}=0$，得$P_s=\frac{2+a_s+6V_s}{9}=\frac{2+a_s}{9}+\frac{2}{3}V_s$

当$P_s<\frac{2+a_s}{9}+\frac{2}{3}V_s$　　$\frac{\partial g}{\partial P_s}>0$

当$P_s>\frac{2+a_s}{9}+\frac{2}{3}V_s$　　$\frac{\partial g}{\partial P_s}<0$

故 $g(P_s)$ 在 $\frac{2+a_s}{9}+\frac{2}{3}V_s$ 处达到最大值。

为了获得线性的 P_s，我们假设[4]

$\frac{1}{3}a_s \le \frac{2+a_s}{9}+\frac{2}{3}V_s \le \frac{1}{3}a_s + 2/3$

且 $g\left(\frac{2+a_s}{9}+\frac{2}{3}V_s\right) \ge 0$， $g\left(\frac{2+a_s}{9}+\frac{2}{3}V_s\right) \ge \frac{1}{3}a_s + \frac{1}{6} - V_s$

这就要求 $3V_s - 2 \le a_s \le 1 + 3V_s$

显然成立。

因 $g\left(\frac{2+a_s}{9}+\frac{2}{3}V_s\right) \ge \frac{\left(\frac{2}{3}a_s + 4/3 - 2V_s\right)^2}{2}$

故 $g\left(\frac{2+a_s}{9}+\frac{2}{3}V_s\right) \ge 0$ 自然成立。

往证 $\frac{\left(\frac{2}{3}a_s + 4/3 - 2V_s\right)^2}{8} \ge \frac{1}{3}a_s + \frac{1}{6} - V_s$

即 $\frac{1}{2}\left(\frac{1}{3}a_s - V_s + 2/3\right)^2 \ge \left(\frac{1}{3}a_s - V_s\right) + \frac{1}{6}$

$\left(\frac{1}{3}a_s - V_s\right)^2 + \frac{4}{3}\left(\frac{1}{3}a_s - V_s\right) + \frac{4}{9} \ge 2\left(\frac{1}{3}a_s - V_s\right) + \frac{1}{3}$

$\left(\frac{1}{3}a_s - V_s\right) - \frac{2}{3}\left(\frac{1}{3}a_s - V_s\right) + \frac{1}{9} \ge 0$

即 $\left(\frac{1}{3}a_s - V_s - \frac{1}{3}\right)^2 \ge 0$

自然成立。

即可。

此时有 $P_b = \frac{1}{3}a_s + \frac{2}{3}V_b$

$$P_s = \frac{2+a_s}{9} + \frac{2}{3}V_s \qquad (6.13)$$

故 $a_s = \frac{2+a_s}{9}$

$c_s = \frac{2}{3}$

$a_b = \frac{1}{3}a_s$

$c_b = \frac{2}{3}$

解得：$a_s = \frac{1}{4}$，$c_s = 2/3$，$a_b = \frac{1}{12}$，$c_b = 2/3$ （6.14）

显然 $a_s = \frac{1}{4}$满足上述约束。

解得：$a_s = \frac{1}{4}$，$a_b = \frac{1}{12}$，$c_b = \frac{2}{3}, c_s = \frac{2}{3}$ ， （6.15）

故线性均衡为：

$$P_b(v_b) = \frac{2}{3}v_b + \frac{1}{12}$$

$$P_s(v_s) = \frac{2}{3}v_s + \frac{1}{4} \quad （6.16）$$

这个博弈的规则要求当且仅当$P_b \geq P_s$才达成交易。

图 6.2 中给出了满足这一条件的区域。

由式（6.16）知交易条件为

$$\frac{2}{3}v_b + \frac{1}{12} \geq \frac{2}{3}v_s + \frac{1}{4}$$

即 $v_b \geq v_s + \frac{1}{4}$

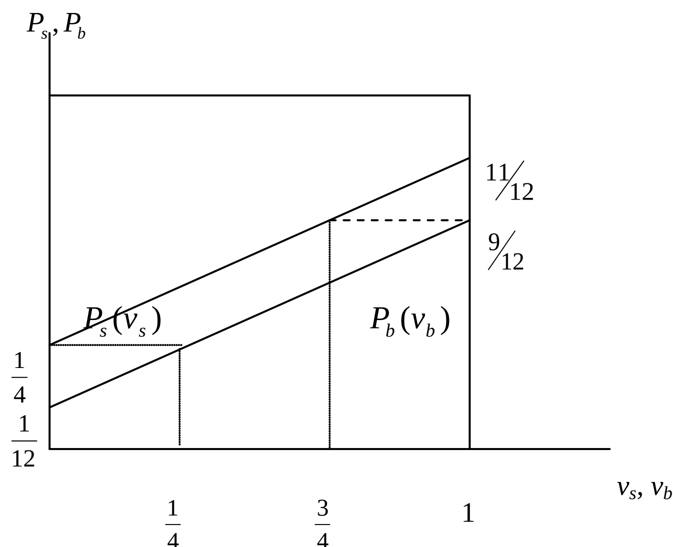


图 6.2　线性交易条件

图 6.3 用更直观的方法给出了交易区域。

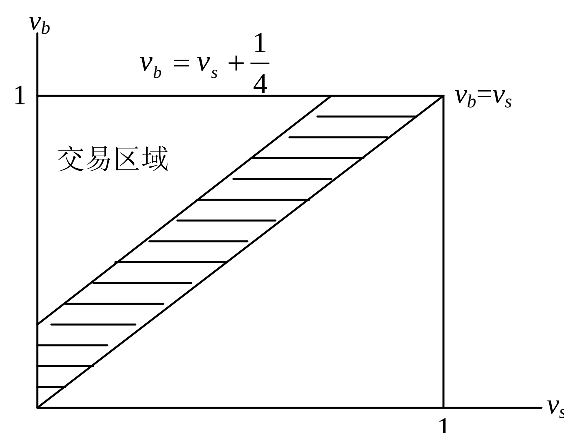


图 6.3　线性交易区域

比较图 6.1 和图 6.3，它们分别表示在单一价格均衡及线性均衡时，交易发生所要求的估价组合范围。在这两种情形、交易的潜在价值最大时即 $v_s = 0$ 且 $v_b = 1$ 时，交易都会达成。但单一价格均衡漏掉了一些有效率的交易如 $v_s = 0$ 且 $v_b = x - \varepsilon$，其中 ε 是充分小的正数，而且还包含了一些基本是无效率的交易如 $v_s = x - \varepsilon$ 且 $v_b = x + \varepsilon$。相反，在线性均衡中，漏过了所有效率较低的交易，只包含了效率或价值至少在 $1/4$ 以上的交易。

这说明从局中人可得到的期望支付看，线性均衡要优于单一价格均衡。

这里并未回答是否还存在另外的均衡，其中局中人的福利还更好一些。Myerson 与 Satterthwaite（1983）曾证明，对于这里假设的估价的均衡分布而言，线性均衡中局中人的期望支付高于博弈的其他任何贝叶斯均衡（包含但不限于单一价格均衡）。这意味着不存在这样的贝叶斯均衡：交易当且仅当有效率时将会发生（即当且仅当 $v_b \ge v_s$ 时）。他们同时证明后一结果是很普遍的，即若 v_b 在区间 $[x_b, y_b]$ 连续分布，v_s 在区间 $[x_s, y_s]$ 连续分布，其中 $y_s > x_b$ 且 $y_b > x_s$，则买方和卖方之间不存在他们所乐于进行的交易博弈，在其贝叶斯均衡中，当且仅当有效率时交易达成。

作为应用，Hall 与 Lazear（1984）提出的就业模型证明：如果企业有关于工人边际产出的私人信息，工人则掌握了自己机会成本的私人信息，则企业与工人之间不存在他们乐于进行的讨价还价博弈，当且仅当雇佣有效率（即边际产出≥机会成本）时达成雇佣协议。

6.4 拍卖与招投标模型

在拍卖或工程项目的招标投标中，不对称信息是一个关键性的特征因素。当一件古董或名画在索斯比拍卖行进行拍卖时，参加竞价拍卖的潜在买主们每一个在其心目

中对古董或名画都有一个价值评价或估价，这个估价别人是不知道的，是每一个潜在买主的“私人信息”。类似地，当一个地方政府打算在流经城区的一条河上修建一座桥梁时，参加承建的建筑公司会来竞标承包这一工程。对每个公司来说，它都有一个最低标价，当政府支付的承包价低于这一最低标价时，公司不会接受承包合同。这一最低标价是每个公司的“私人信息”，别人是不清楚的。这样，参加拍卖的潜在买主们每人心中有一最高价格，它是每个潜在买主的私人信息，而参加竞标的每个公司都有一个别的公司和政府都不清楚的最低标价，它是每个竞标公司的私人信息。

拍卖的或竞标的在本质上是相同的，以下仅就拍卖的经济机理进行博弈分析，同样的结论完全适用于招投标。

拍卖有两个目的，一是寻找最高出价的买家，二是在直接的卖者或买者以代理人身份出现时，拍卖也有助于减少买者与卖者之间的损害委托人的合谋。譬如，一个城市的市长可以任意出租土地给企业时，很难保证得到土地的企业不是贿赂市长但只付很低租金的企业。但若采用公开拍卖方式出租土地，市长接受贿赂的可能性就小得多。

如果拍卖是买主们公开透明地报出其买价，且其他的买主能知悉任一买主报出的价格，则拍卖就容易出现“围标”现象。所谓“围标”，就是买主们串通起来压低买价（或卖主们抬高卖价）。由于任一买主的报价都是共同知识，所以这种串谋是可以实施的，因为串通起来的买主群体能够对擅自报出较高买价（或较低卖价）的人进行惩罚[4]。

当拍卖一件古董或一幅名画时，“围标”会使卖者蒙受损失。譬如，古董价值在买主们的估价中最高为 100 万元，但围标可能使它最终以 90 万元卖出去，因为围标者达成的协议是任何买主报出的买价最多不超过 90 万元。

解决围标的一种办法是通过减低买主们报价的透明性。譬如，当买主的估价最高为 100 万元时，如果每个买主报出的价格别的买主都不知道，则估价最高的那位买主要报出高于估价第二高的估价但低于 100 万元就可获得古董或名画，同时还有正的消费者剩余。当然，这里潜在假定拍卖规则是如此规定的：每一位买主独立地（别的买主不知道地）报出其买价，卖主根据报价最高的那位买主报出的价格将古董或名画卖给报价最高的人。当多位不同的买主报出相同的价格时，由他们之间按掷骰子的方式随机决定最终获得古董或名画的买主，但支付的价格仍为他们早先报出的最高价格。这就是所谓“一级密封价格拍卖（the first—price Sealed auction）。这种拍卖方法是要求买主们同时将自己的出价写下来装入一个信封并密封后交给拍卖人。拍卖人打开信封，决定将古董或名画卖给出价最高的买主，按他的出价支付价格。当多位买主同时报出相同最高价格时，拍卖者通过掷骰子随机地将古董或名画卖给其中的一位买主，按他们报出的最高价格支付价格。一级密封价格拍卖可以避免围标，因为每一个买主的报价对于其他买主来说是不可观测的（当然，要做到这样，还要求成交后拍卖者要

对买主的报价保密），因而每个买主的行为不受围标组织者的约束。一级密封价格拍卖尽管可以避免围标，但不能保证买主们会报出他们心目中的真实估价或最高价格，即关于古董或名画值多少钱买主们不会说真话。譬如，当最高估价者只有一位且其估价为 100 万元，则估价最高者知道，当他报出 100 万元的价格时，尽管可能拿到古董或名画，但其消费者剩余为零，而当他的报价低于 100 万元时，他有正的概率获得商品，此时有正的消费者剩余。因而其预期消费者剩余即消费者剩余的期望值为正。我们假定买主们是风险中性的，则他的最优选择是报出低于 100 万元但大于零的价格。其实，对于那些估价不是最高的 100 万元价格的买主来说，同样的道理使他们相信最优选择是报出低于他们心中的估价但大于零的价格。关于这一点，我们下面给出一个模型加以说明。设有一件古董需要按一级密封价格拍卖进行拍卖，有 n 个完全相同的（支付函数完全相同）的买主参加，第 i 位买主的支付函数为其消费者剩余：

$$u_i(b_i,b_j,v_i)=\begin{cases}v_i-b_i,\ b_i>b_j,\ j\neq i\\ \dfrac{1}{m}(v_i-b_i),\ \text{若有}m\leq n\text{位买主的报价}b_k(\text{包括第}i\text{位买主})\text{与}b_i\text{相等且}b_i\geq b_j,\ j\neq i\\ 0,\ \mathbf{b_i}<\text{某个}b_j,\ i\neq j\end{cases}\quad(6.17)$$

其中，b_i 是第 i 位买主的报价。

v_i 是第 i 位买主的估价，假定对于第 j 位买主来说 $(j\neq i)$， v_i 是区间 $[0,1]$ 上的均匀分布的贝叶斯随机变量。

由于多位买主同时报出相同价格的概率在连续分布情形实际上为零，所以对此情形不予考虑在数学上会方便一些。设 $b_i(v_i)$ 是类型依赖的均衡报价战略，不妨设 $b_i(v_i)$ 是严格增函数。显然，$b_i>1$ 不是最优的，因为此时的预期支付为负，还不如 $b_i=v_i$ 带来的预期支付（为零）大。由于这是一个对称博弈，我们下面只考虑对称的均衡出价战略，即 $b_i(v_i)=b_j(v_j)$，当 $v_i=v_j$。令均衡出价战略的一般形式为

$b=b^*(v)$

给定 v，第 i 个买主选择报价 b 的期望或预期支付为

$$u_i=(v-b)prob(b>b_j,j\neq i)\qquad（6.18）$$

其中 $prob(b>b_j,j\neq i)$ 是第 i 位买主报价最高的概率。

这里的一级密封价格拍卖实际上保证了各个买主的报价是相互独立的，因而有 $prob(b>b_j,j\neq i)=\prod_{j\neq i}prob(b>b_j)$。

令 $\Phi(b)=b^{*-1}(b)$，则

$$
\begin{aligned}
prob(b > b_j) &= prob(b^*(v_j) < b) \\
&= prob(v_j < b^{*-1}(b)) \\
&= prob(v_j < \Phi(b)) \\
&= \Phi(b)
\end{aligned}
$$

最后一个等式由均匀分布的假定得出。

于是，第 i 位买主的最大化问题是

$$\max_b \left\{ u_i = (v-b)\prod_{j\neq i} prob(b_j < b) = (v-b)\Phi^{n-1}(b) \right\}$$

一阶条件

$$-\Phi^{n-1}(b) + (n-1)(v-b)\Phi^{n-2}(b)\Phi'(b) = 0 \qquad (6.19)$$

上式两端同除以 $\Phi^{n-2}(b)$：

$$-\Phi(b) + (n-1)(v-b)\Phi'(b) = 0$$

或

$$-v + (n-1)(v-b)\frac{dv}{db} = 0$$

解微分方程得 $b^*(v) = \dfrac{n-1}{n}v < v$，每个买主都低报价格。

显然，$b^*(v)$ 随 n 的增加而增加，当 n 趋向无限大时，b^* 趋于 v，即买主愈多，卖者得到的价格就愈高，故让更多的人加入竞标在招投标中是符合招标人的利益的。在公共管理中，政府的采购和公共工程招投标中通常规定要进行公开招标，并在参加竞标的公司数目上有下限规定，其缘故正是如此，因为更多的竞争者参加投标会压低工程报价，从而使政府开支得到一定程度的节省。

美国经济学家 Vickrey 发明的“二级密封价格拍卖”（The Second-Price Sealed auction，又称为 Vickrey 招投标法），可以使买主们报出真实的估价，即让买主们说出真话。这一种方法目前已是国际性招投标中常用的方法，但在 1970 年代，当时美国联邦政府为大量公共工程招投标中严重的围标现象感到头疼时，经济学家们向他们推荐这种“Vickrey 招投标法”。由此，联邦政府在公共工程中随后节省了大量开支。Vickrey 于 1996 年与 Mirrles 一道获得当年的经济学诺贝尔奖，Vickrey 招投标法是一个重要因素，因为它是以往被认为是过于抽象的博弈论走向实用的一个重要代表。

与一级密封价格拍卖法唯一不同的是，二级密封价格拍卖是将拍卖品卖给出价最高的买主的同时要求其支付第二高的报价。此时，若他赢得古董或名画，则他支付的价格或成本对于他的各种可能的报价选择来说是固定的，即他的报价选择与其潜在的成本无关。这样，给定成本不变，买主会极大化其赢得拍卖品的概率，这就使得他会尽量选择高的报价。买主报出的价格不会超过他的估价，这是因为，与报出等于估价的价格相比，报出高于估价的报价是劣战略。事实上，当报价高于估价时，若赢得拍

卖品，则当第二高报价高于他自己的估价时，他的支付是负的，小于当他报出等于估价的支付（为零），当第二高报价低于他自己的估价时，他报出等于自己估价的价格也会赢得拍卖品且带来相同的支付。与估价相比，报价低于估价的战略也是劣战略。因为，当他报价低于估价时，他赢得拍卖品的概率小于报出估价的情形，而当报价不超过估价时，一旦赢得拍卖品，支付是相同的。下面在均匀分布下证明这一结论。

当第 1 位买主选择报价为 $b_1(v_1)$ 时，则期望支付为

$$u_1=\sum_{\substack{j=2\\k=2\\j\neq k}}\int_{b_1(v_1)>b_j>b_k}[v_1-b_j(v_j)]P(v_2,\cdots,v_n)dv_2\cdots dv_n$$

因为任何两个买主报相同价格的概率在连续分布下为零，故这里不考虑有多个买主报相同价格的情形。

假定 $b_j(v_j)$ 是严格增函数，则由对称性及独立性有

$$u_1=\sum_{j=2}^{n}\int_0^{b_j^{-1}(b_1)}dv_j\underbrace{\int_0^{b_2^{-1}(b_j)}\cdots\int_0^{b_k^{-1}(b_j)}\cdots\int_0^{b_n^{-1}(b_j)}}_{k\neq j,1}[v_1-b_j(v_j)]P(v_2\cdots v_n)dv_2\cdots dv_k\cdots dv_n$$

$$=\sum_{j=2}^{n}\int_0^{b_j^{-1}(b_1)}[v_1-b_j(v_j)]P(v_j)dv_j\underbrace{\int_0^{b_2^{-1}(b_j)}P(v_2)dv_2\cdots\int_0^{b_k^{-1}(b_j)}P(v_k)dv_k\cdots\int_0^{b_n^{-1}(b_j)}P(v_n)dv_n}_{k\neq j,1}$$

$$=\sum_{j=2}^{n}\int_0^{b_j^{-1}(b_1)}[v_1-b_j(v_j)]\underbrace{b_2^{-1}(b_j)\cdots b_k^{-1}(b_j)\cdots b_n^{-1}}_{k\neq j,1}(b_j)dv_j$$

当 $v_1\geq b_j(v_j)$ 时，它显然是 b_1 的严格增函数，这要求 $b_j(b_1^{-1}(b_1))=b_1\leq v_1$，故最优的 b_1 为 v_1。由对称性有 $b_j=v_j, j=1,\cdots,n$。这样，二级密封价格拍卖使买主报价他们心目中对拍卖品的真实估价。如果说每个买主的估价相同，则这种方法将拍卖品以最高的价格卖出去，卖者获得最大的拍卖价值。当不同买主的估计不同但相差不大时，卖者也会获得较大的拍卖价值。这正是二级密封价格拍卖的优点，它既避免了围标，也以较大的价值将拍卖品卖出或以较低的价格招标。这种方法的一个重要的功能是诱使买主或投标人说出真话。这正是博弈论的机制设计功能，我们将在后面系统加以介绍。

6.5 最优拍卖机制设计：Myerson 定理

在 6.4 节中，我们介绍了几种拍卖或招投标方法。一个顺理成章的问题是，在不同的拍卖或招投标方法中，有没有一种方法对于拍卖人或招标人来说是最优的。这是

因为，尽管二级密封招标法优于一级密封招标法，但或许还存在更好的招标法呢。Myerson（1981）在其经典论文中对此进行了研究，本节将概要地介绍他的工作。

这个问题的一般性表述是：

一个卖者打算将其拥有的一件物品卖给 n 位打算购买这一物品的潜在购买者中的某一位，但卖者不知道这些潜在的购买者所愿意为这一物品所支付的最高价格，即卖者对于不同买者对于该物品的价值评价或支付意愿具有不完全信息。现在，卖者的问题是如何设计一个拍卖博弈规则，它存在一个纳什均衡，该均衡为卖者带来一个最高的期望支付（或期望效用）。下面，我们遵循 Myerson 的方法来导出相当广泛意义上的一大类可能的拍卖机制设计中的最优拍卖机制。

6.5.1 基本假定与定义

假定有一位拍卖者打算把其拥有的一件物品出售。他面对 n 个潜在的购买人，记为 $1,2,\cdots,n$。用 N 表示潜在买者构成的集合，即

$$N=\{1,\cdots,n\} \tag{6.20}$$

记第 i 个买者对物品的价值评价或支付意愿为 v_i，它是买者的私人信息，$i\in N$。假设对于卖者和其他买者来说，他们对买者 i 的价值评价 v_i 的信念是：v_i 是 $[a_i,b_i]$ 上的随机变量（贝叶斯统计学意义上），且密度函数为 $f_i(v_i)$，并且这也是包括所有买者和卖者在内的局中人的共同知识，假定 $-\infty<a_i<b_i<+\infty$，$f_i(v_i)>0$，$\forall v_i\in[a_i,b_i]$；且 $f_i(v_i)$ 是 $[a_i,b_i]$ 上的连续函数。设 v_i 的分布函数为 $F_i(v_i)$，即

$$F_i(v_i)=\int_{a_i}^{v_i} f_i(s_i)ds_i \tag{6.21}$$

记 V 为买者们价值评价的所有可能的组合：

$$V=\prod_{i\in N}[a_i,b_i] \tag{6.22}$$

记 V_{-i} 为除买者 i 之外的其他买者的价值评价组合，即

$$V_{-i}=\prod_{\substack{j\in N\\ j\neq i}}[a_j,b_j] \tag{6.23}$$

假定 n 个买者的价值评价是互相独立的随机变量，于是 V 上的联合密度函数 $f(v)$ 满足如下关系：

$$f(v)=\prod_{j\in N} f_j(v_j) \tag{6.24}$$

其中 $v=(v_1,\cdots,v_n)$。

假定买者 i 和卖者在 v_{-i} 上的联合密度函数为：

$$f_{-i}(v_{-i})=\prod_{\substack{j\in N\\ j\neq i}} f_j(v_j) \tag{6.25}$$

其中，$v_{-i}\in V_{-i}$。

记卖者自己对物品的估价为v_0，并设它是所有博弈局中人的共同知识。局中人对物品的估价之所以是私人信息，存在两个方面的原因。首先，某局中人的偏好是其他局中人不知道的。譬如，当物品是一幅画时，其他人并不知道他在欣赏这幅画的过程中所获得的愉悦程度。其次，买者可能拥有一些关于物品内在质量（intrinsic quality）的特殊信息。我们分别称这两种因素为偏好的不确定性（preference uncertainty）和质量不确定性（quality uncertainty）。这种区分是十分重要的。如果仅存在偏好不确定性，将买者j的估价信息告诉买者i将不会引起买者i修正其估价（这并不意味着当买者i知道买者j的估价后不会修改其报价战略，这仅意味着买者i在拥有货币与物品之间的原初偏好顺序是不变的）。然而，当存在质量不确定时，买者在知道了其他人对物品的估价后将倾向于修正其自己的估价。也就是说，如果买者i知道v_j很低的话，他会认为买者j收到了有关物品质量的不好的信息，于是，买者i将根据自己的判断也降低自己对物品的估价。在许多讨论拍卖的文献中，仅就只存在偏好不确定性的情形加以考虑，譬如，Vickrey（1961）。与之不同，Myerson 拓宽了这一研究的范围。他的研究将一定形式的质量不确定性因素也考虑在内。假定存在n个“修正效应函数”（revision effect function）e_j，其定义如下：

e_j是定义在$[a_i, b_i]$上的实函数，如果局中人i知道了v_j，则i会通过$e_j(v_j)$对自己的估价加以修正[5]。因此，若i知道了$v = (v_1, \cdots, v_n)$是n个买者最初的估价向量，则i将修正其估价如下：

$$\tilde{v}_i(v) = v_i + \sum_{\substack{j \in N \\ j \neq i}} e_j(v_j) \tag{6.26}$$

类似地，我们假定卖者也会重新对物品进行估价：

$$\tilde{v}_0(v) = v_0 + \sum_{j \in N} e_j(v_j) \tag{6.27}$$

在仅存在纯偏好不确定性假设下，$e_j(v_j) \equiv 0$[6]。

6.5.2 可行的拍卖机制

给定上述密度函数f_i，修正函数e_i和$\tilde{v}_i$，卖者的问题是选择一种拍卖机制最大化其期望支付（或期望效用）。我们只需将注意力限制在一类特别的拍卖机制上，这就是“直接显示机制”（direct revelation mechanism）。

在直接显示机制中，买者们同时且老老实实地向卖者揭示其估价，卖者决定谁将会买得物品和买者需要支付的价格，这一价格是买者们报出的估价向量$v = (v_1, \cdots, v_n)$的某种函数。因此，直接显示机制可以表达为一对产出函数(P, x)（定义为$P: V \to R^n$和$x: V \to R^n$），满足如下性质：如果v是揭示出来的估价向量，则$P_i(v)$是i得到物品的概率和$x_i(v)$是i必须支付给卖者的货币量[7]。

以下总假定卖者和买者都是风险中性的（risk neutral），并且对于货币和物品具有可加的可分性效用函数（additively separable utility functions）。因此，给定 i 知道 v_i，其在一个特定的拍卖机制 (P,x) 下的期望支付为

$$U_i(P,x,v_i)=\int_{v_{-i}}[\tilde{v}_i(v)P_i(v)-x_i(v)]f_{-i}(v_{-i})dv_{-i} \tag{6.28}$$

其中 $dv_{-i}=dv_1\cdots dv_{i-1}dv_{i+1}\cdots dv_n$。

类似地，卖者在给定的这一拍卖机制中获得的期望支付为

$$U_0(P,x)=\int_v\left[\tilde{v}_0(v)\left(1-\sum_{j\in N}P_j(v)\right)+\sum_{j\in N}x_j(v)\right]f(v)dv \tag{6.29}$$

其中 $dv=dv_1\cdots dv_n$。

并非每一对函数 (P,x) 都表达一种可行的拍卖机制，然而，有三种类型的约束必须施加于 (P,x)。

首先，因为仅有一种待售物品，函数 P 须满足下列概率条件：

$$\sum_{j\in N}P_j(v)\le 1 \text{且} P_j(v)\ge 0，\ \forall i\in N，\ \forall v\in V \tag{6.30}$$

其次，当潜在买者不参加拍卖时，其期望效用为零，故欲使买者参与拍卖，就有如下的“个人理性”（individual rationality）约束条件：

$$U_i(P,x,v_i)\ge 0，\ \forall i\in N，\ \forall v_i\in[a_i,b_i] \tag{6.31}$$

最后，假定卖者并不能使买者在其估价上对他不说谎话。

因此，只有没有一位买者会从说谎中获益，直接显示机制才可能被实施。这就是说，此时说实话才会是一个纳什均衡。如果买者 i 在其真实估价为 v_i 时宣称其估价为 s_i，则期望支付为

$$\int_{v_{-i}}\left[\tilde{v}_i(v)P_i(v_{-i},s_i)-x_i(v_{-i},s_i)\right]f_{-i}(v_{-i})dv_{-i}$$

其中 $(v_{-i},s_i)=(v_1,\cdots,v_{i-1},s_i,v_{i+1},\cdots,v_n)$。因此，为了保证没有一位买者有谎报其估价的动机，下列“激励相容”（incentive compatibility）约束必成立。

$$U_i(P,x,v_i)\ge\int_{v_{-i}}\left[\tilde{v}_i(v)P_i(v_{-i},s_i)-x_i(v_{-i},s_i)\right]f_{-i}(v_{-i})dv_{-i} \tag{6.32}$$

$$\forall i\in N,\forall v_i\in[a_i,b_i],\ \forall s_i\in[a_i,b_i]$$

当且仅当式（6.32）、（6.31）、（6.30）同时成立时，称机制 (P,x) 是可行的（feasible）（或称机制 (P,x) 表达了一种可行的拍卖机制）。也就是说，若卖者打算据 P 配置物品和据 x 索取货币支付，则当且仅当式（6.30）—（6.32）成立时，该程序将在买者们说真话的情况下得到贯彻。

到目前为止，我们仅考虑了其中买者们总是说真话的直接显示机制。但是，卖者也可设计其它类型的拍卖博弈。在一般性的拍卖博弈中，每个买者拥有某些备选战略集 Θ_i，并且存在产出函数

$\hat{P}:\prod_{i\in N}\Theta_i \to R^n$ 且 $\hat{x}:\prod_{i\in N}\Theta_i \to R^n$

它描述了物品配置及买者支付的费用是如何依赖于买者们的战略的（即若 $\theta=(\theta_1,\cdots,\theta_n)$ 是买者们选择的战略组合，则 $\hat{P}_i(\theta)$ 就是 i 获得物品的概率和 $\hat{x}_i(\theta)$ 是 i 向卖者支付的期望支付）。

一个拍卖机制就是这样的一种拍卖博弈且附加了关于买者们打算选择的战略计划的一种描述。形式上，一个战略计划（Strategic Plan）可表达为一个函数 $\hat{\theta}_i:[a_i,b_i]\to\Theta_i$ 满足当 i 的估价为 v_i 时，$\hat{\theta}_i(v_i)$ 是 i 打算选择的战略。在此一般性的规定下，直接显示机制就显然是这样一种拍卖机制即 $\Theta_i=[a_i,b_i]$ 和 $\hat{\theta}_i(v_i)\equiv v_i$。

在这个一般性的框架内，一个可行的拍卖机制必须满足式（6.30）—（6.32）给出的约束条件。由于仅有一件待售品。对于任何 θ 来说，概率 $\hat{P}_i(\theta)$ 必须为非负的且总和不会大于 1。给定任何估价值，拍卖机制必须对每个买者都给出非负的期望支付，否则买者不会参与拍卖。在一个拍卖博弈中，所有的战略计划必须构成一个纳什均衡，否则某些买者将修正其计划。

显而易见的是，最优拍卖设计问题是难以求解的，因为对于卖者在其构造的拍卖博弈中所要使用的战略空间 Θ_i 的大小和复杂程度没有任何限制。使我们能解决拍卖机制设计问题的一个基本的出发点是我们仅考虑直接显示机制时并未损失一般性。这一结论来自如下事实：

引理 6.1（显示性原理）给定任何一个可行拍卖机制，存在一个等价的可行的直接显示机制（equivalent feasible direct revelation mechanism），它给予卖者及所有的买者予与给定机制完全相同的期望支付。

Myerson（1979）曾在更为一般的贝叶斯公共选择问题中证明了这个显示性原理，但这里我们可以根据直观明白该引理成立的原因。假定给定一个可行的拍卖机制，其战略空间为 Θ_i，产出函数为 $\hat{P}$ 和 $\hat{x}$；战略计划为 $\hat{\theta}_i$。考虑如下定义的一个直接显示机制 (P,x)，其中 $P:V\to R^n$； $x:V\to R^n$，且

$$P(v_1,\cdots,v_n)=\hat{P}(\hat{\theta}_1(v_1),\cdots,\hat{\theta}_n(v_n))$$

$$x(v_1,\cdots,v_n)=\hat{x}(\hat{\theta}_1(v_1),\cdots,\hat{\theta}_n(v_n))$$

也就是说，在直接显示机制 (P,x) 中，卖者首先要求每一个买者说出其类型（即估价），然后根据给定的拍卖机制的战略计划计算出买者将选择的战略。最后实施由这些战略给出的产出。因此，直接显示机制 (P,x) 总给出与给定拍卖机制相同的产出，故所有局中人在两种机制中获得的期望支付相同。(P,x) 必满足式（6.32）给出的激励相容约束，因为给定可行机制中的战略计划构成了一个均衡（如果任何一位买者在显示性博弈中可以通过向卖者撒谎来获取好处，则他可以在给定机制中通过“向自己撒谎”或修正他的战略计划来获益）。因此，(P,x) 是可行的。

采用显示性原理，我们不失一般性地假设卖者只考虑可行的直接显示机制中的拍卖机制。也就是说，我们今后将可行的拍卖机制与所有满足式（6.30）、（6.31）、(6.32) 的产出函数 (P,x) 的集合等同起来。卖者的拍卖设计问题就是选择这样的函数 $P:V\to R^n$ 和 $x:V\to R^n$ 使得 $U_0(P,x)$ 在式（6.30）—（6.32）约束下达到最大化。

注意，我们在这里并未用到式（6.25）或式（6.26）。所以，式（6.30）—（6.32）刻画了所有可行拍卖机制的特征，甚至包括买者用函数 $\hat{v}_i:V\to R$ 计算其修正估价 $\hat{v}_i(v)$ 的情形，而这种函数可以不是式（6.26）给出的那种特殊的可加形式。然而，当 Myerson 为导出最优拍卖设计的显性解时，他不得不将研究限于式（6.25）及(6.26) 给出的范围之内。

6.5.3 模型分析

给定拍卖机制 (P,x)，对于任意的买者 i 和任意的估价 v_i，定义

$$\theta_i(P,v_i)=\int_{v_{-i}}P_i(v)f_{-i}(v_{-i})dv_{-i} \qquad (6.33)$$

$\theta_i(P,v_i)$ 是买者 i 在给定其估价为 v_i 时将在拍卖机制 (P,x) 中获得物品的条件概率。

我们下面首先要简化对可行拍卖机制的刻画。于是有如下引理：

引理 6.2　(P,x) 是可行的当且仅当下列条件成立：

若 $s_i\le v_i$ 则 $\theta_i(P,s_i)\le\theta_i(P,v_i),\ \forall i\in N,\forall s_i,v_i\in[a_i,b_i]$ （6.34）

$$U_i(P,x,v_i)=U_i(P,x,a_i)+\int_{a_i}^{v_i}\theta_i(P,s_i)ds_i,\forall i\in N,\forall v_i\in[a_i,b_i] \qquad (6.35)$$

$$U_i(P,x,a_i)\ge 0,\forall i\in N \qquad (6.36)$$

和

$$\sum_{j\in N}P_j(v)\le 1 \text{ 且 } P_i(v)\ge 0，\ \forall i\in N,\forall v\in V \qquad (6.37)$$

证明：利用式（6.26）即我们关心 $\tilde{v}_i(v)$ 的函数形式的一个特别的假定，有

$$\tilde{v}_i(v)-\tilde{v}_i(v_{-i},s_i)=v_i-s_i$$

$$\text{即}\tilde{v}_i(v)=\tilde{v}_i(v_{-i},v_i)$$
$$=\tilde{v}_i(v_{-i},s_i)+v_i-s_i$$

$$\int_{v_{-i}}[\tilde{v}_i(v)P_i(v_{-i},s_i)-x_i(v_{-i},s_i)]f_{-i}(v_{-i})dv_{-i}$$
$$=\int_{v_{-i}}[[\ (\tilde{v}_i(v_{-i},s_i)+(v_i-s_i)]P_i(v_{-i},s_i)-x_i(v_{-i},s_i)]\ ]f_{-i}(v_{-i})dv_{-i}$$
$$=U_i(P,x,s_i)+(v_i-s_i)\theta_i(P,s_i)$$

所以，激励相容约束式（6.32）等价于

$$U_i(P,x,v_i) \ge U_i(P,x,s_i) + (v_i - s_i)\theta_i(P,s_i), \forall i \in N, \forall v_i, s_i \in [a_i, b_i] \qquad (6.38)$$

故(P,x)是可行的当且仅当式（6.30）、（6.31）和式（6.38）成立。我们现在将证明式（6.31）及（6.38）意味着式（6.34）—（6.36）成立。两次运用式（6.38）（将s_i与v_i互换），有

$$(v_i - s_i)\theta(P,s_i) \le U_i(P,x,v_i) - U_i(P,x,s_i) \le (v_i - s_i)\theta_i(P,v_i)$$

则式（6.34）在$s_i \le v_i$时成立。

这些不等式在对于任意的$\delta > 0$都可重写为：

$$\theta_i(P,s_i)\delta \le U_i(P,x,s_i+\delta) - U_i(P,x,s_i) \le \theta_i(P,s_i+\delta)\delta$$

令$\delta \to 0$,则

$$\theta_i(P,s_i) \le \frac{\partial U_i}{\partial s_i} \le \theta_i(P,s_i)$$

因$\theta_i(P,s_i)$对于s_i是递增的，它是黎曼可积的，故

$$\int \theta_i(P,s_i)ds_i \le \int \frac{\partial U_i}{\partial s_i}ds_i \le \int \theta_i(P,s_i)ds_i$$

$$\int_{a_i}^{v_i} \theta_i(P,s_i)ds_i = U_i(P,x,v_i) - U_i(P,x,a_i)$$

它给出式（6.35）。

当然，式（6.36）直接由式（6.31）导出，式（6.37）就是式（6.30），故引理 2 中的所有条件都从可行性得出。

现在我们必须证明引理 2 中的条件也意味着式(6.31)和式(6.38)（从而式(6.32)）成立。

因据式(6.37)有$\theta_i(P,s_i) \ge 0$，而式(6.31)来自式(6.35)和(6.36)及$\theta_i(P,s_i) \ge 0$。

为了证明式（6.38），假设$s_i \le v_i$，则由式（6.34）和（6.35）有：

$$\begin{aligned} U_i(P,x,v_i) &= U_i(P,x,s_i) + \int_{s_i}^{v_i} \theta_i(P,v_i)dv_i \\ &\ge U_i(P,x,s_i) + \int_{s_i}^{v_i} \theta_i(P,s_i)dv_i \\ &= U_i(P,x,s_i) + (v_i - s_i)\theta_i(P,s_i) \end{aligned}$$

类似地，若$s_i > v_i$则

$$\begin{aligned} U_i(P,x,v_i) &= U_i(P,x,s_i) - \int_{v_i}^{s_i} \theta_i(P,v_i)dv_i \\ &\ge U_i(P,x,s_i) - \int_{v_i}^{s_i} \theta_i(P,s_i)dv_i \\ &= U_i(P,x,s_i) + (v_i - s_i)\theta_i(P,s_i) \end{aligned}$$

故式（6.38）由式（6.34）和（6.35）导出。当然，式（6.30）就是式（6.37）。所以，由引理 2 中的条件也导出可行性。证毕！

所以，(P,x) 是一种最优拍卖当且仅当它在约束（6.35）—（6.37）和式（6.34）下最大化 $U_0(P,x)$。下面的一个引理为最优性提供了某些简单的条件。

引理 6.3　假设 $P:V\to R^n$ 在约束（6.30）和（6.34）下最大化如下积分

$$\int_v\left[\sum_{i\in N}[v_i-e_i(v_i)-\frac{1-F_i(v_i)}{f_i(v_i)}-v_0]P_i(v)\right]f(v)dv \tag{6.39}$$

还假定

$$x_i(v)=P_i(v)\tilde{v}_i(v)-\int_{a_i}^{v_i}P_i(v_{-i},s_i)ds_i,\forall i\in N,\forall v\in V \tag{6.40}$$

则 (P,x) 就是一种最优拍卖。

证明：回忆式（6.29），我们可将卖者的目标函数改写为：

$$\begin{aligned}U_0(P,x)&=\int_v\tilde{v}_0(v)f(v)dv+\sum_{i\in N}\int_v P_i(v)[\tilde{v}_i(v)-\tilde{v}_0(v)]f(v)dv\\&\quad+\sum_{i\in N}\int_v[x_i(v)-P_i(v)\tilde{v}_i(v)]f(v)dv\end{aligned} \tag{6.41}$$

但运用引理 6.2，可知对任何可行的 (P,x) 有：

$$\begin{aligned}&\int_v[x_i(v)-P_i(v)\tilde{v}_i(v)]f(v)dv\\&=-\int_{a_i}^{b_i}U_i(P,x,v_i)f_i(v_i)dv_i\\&=-\int_{a_i}^{b_i}[U_i(P,x,a_i)+\int_{a_i}^{v_i}\theta_i(P,s_i)ds_i]f_i(v_i)dv_i\\&=-U_i(P,x,a_i)-\int_{a_i}^{b_i}\int_{a_i}^{v_i}f_i(v_i)\theta_i(P,s_i)dv_ids_i\\&=-U_i(P,x,a_i)-\int_{a_i}^{b_i}\int_{a_i}^{v_i}\theta_{ii}(P,s_i)ds_idF_i(v_i)\\&=-U_i(P,x,a_i)-\left[\int_{a_i}^{b_i}\theta_i(P,s_i)ds_i-0-\int_{a_i}^{b_i}F_i(v_i)\theta_i(P,v_i)dv_i\right]\\&=-U_i(P,x,a_i)-\int_{a_i}^{b_i}[1-F_i(v_i)]\theta_i(P,v_i)dv_i\\&=-U_i(P,x,a_i)-\int_v[1-F_i(v_i)]P_i(v)f_{-i}(v_{-i})dv\end{aligned} \tag{6.42}$$

由式（6.25）和（6.26）有

$$\tilde{v}_i(v)-\tilde{v}_0(v)=v_i-v_0-e_i(v_i) \tag{6.43}$$

将式（6.42）和（6.43）代入式（6.41），得

$$\begin{aligned}U_0(P,x)&=\int_v\left[\sum_{i\in N}\left[v_i-v_0-e_i(v_i)-\frac{1-F_i(v_i)}{f_i(v_i)}\right]P_i(v)\right]f(v)dv\\&\quad+\int_v\tilde{v}_0(v)f(v)dv-\sum_{i\in N}U_i(P,x,a_i)\end{aligned} \tag{6.44}$$

故卖者的问题就是在引理 2 中的约束（6.35）、（6.36）、（6.37）和（6.34）下最大化式（6.44）。在此表述中，x 仅出现在目标函数中的最后一项和约束（6.35）、（6.36）

中。这两个约束可重写成：

$$\int_{v_{-i}}[P_i(v)\widetilde{v}_i(v)-\int_{a_i}^{v_i}P_i(v_{-i},s_i)ds_i-x_i(v)]f_{-i}(v_{-i})dv_{-i}$$

$$=U_i(P,x,a_i)\geq 0,\ \forall i\in N,\forall v_i\in[a_i,b_i]$$

若卖者据式（6.40）选择x，则他同时满足式（6.35）和（6.36）且他获得

$$\sum_{i\in N}U_i(P,x,a_i)=0$$

它是式（6.44）中该项的可能最好的取值。

所以，利用式（6.40），我们可以在卖者的问题中整个将x丢掉。进一步，式（6.44）右端中的第二项是独立于(P,x)的常数。所以，目标函数可被简化为式（6.39），且式（6.34）和（6.37）是唯一需要满足的约束条件。

证毕！

方程（6.44）有一个重要推论，它是如此重要以至于下面将其作为一个定理给出

定理 6.1 （收益等价性定理）卖者从一个可行拍卖机制中获得的期望支付完全由概率函数P及$U_i(P,x,a_i)$（对所有i）所决定。

也就是说，一旦我们知道了在每种可能情况下（由P刻画）谁获得了物品，以及每个买者在其估价为其最低可能水平a_i时的期望支付为多少，则卖者从拍卖中得到的期望支付并不依赖于产出函数x。因此，譬如，对于任意两个拍卖机制，只要它们具有性质（1）物品总是卖给高于v_0的最高估价者和（2）如果其估价位于其最低可能估价,则买者的期望支付为零，则卖者会从它们中获得相同的期望支付。

6.5.4 正规情况下的最优拍卖

在一种简单的通常假定下，我们能够直接从引理 6.3 计算出最优拍卖机制。如果函数

$$c_i(v_i)=v_i-e_i(v_i)-\frac{1-F_i(v_i)}{f_i(v_i)} \qquad (6.45)$$

是v_i的严格递增函数（对N中的每一个i），我们称问题是正规的（regular）。也就是说，当$a_i\leq s_i<v_i\leq b_i$，若$c_i(s_i)<c_i(v_i)$，则问题是正规的（我们曾假定$f_i(v_i)>0$(对所有$[a_i,b_i]$中的v_i)，故$c_i(v_i)$总有定义且是连续的）。

现在，考虑一种拍卖机制，其中当$v_0>\max_{i\in N}[c_i(v_i)]$时卖主不出售物品，否则他把物品卖给报出最高$c_i(v_i)$的买主。如果$c_i(v_i)=c_j(v_j)=\max_{k\in N}[c_k(v_k)]\geq v_0$，则卖主可以通过将物品交给最低编号的买主或其它的随机性规则解决配置问题(在正规情况下这种情形发生的概率为零)。所以，对于这一拍卖机制，有

$$P_i(v)>0\text{ 意味着 }c_i(v_i)=\max_{j\in N}[c_j(v_j)]\geq v_0 \qquad (6.46)$$

对所有V中的v，这一机制最大化如下的和

$$\sum_{i\in N}[c_i(v_i)-v_0]P_i(v)$$

约束条件为

$$\sum_{j\in N}P_j(v)\le 1 \text{且} P_i(v)\ge 0, \quad \forall i$$

故P在约束条件式（6.30）下最大化式（6.39）。为了验证它也满足式（6.34），我们需要使用正规性假设。假设$s_i<v_i$，$c_i(s_i)<c_i(v_i)$，且当买主i提交一个估价s_i就会赢得物品时，如果他将估价变为v_i，他也会赢。也即是对所有的v_{-i}有$P_i(v_{-i},s_i)\le P_i(v_{-i},v_i)$。故在$v_i$为给定估价时，$i$赢得物品的概率$\theta_i(P,v_i)$是$v_i$的增函数，这正是式（6.39）所需的。所以$P$满足引理 6.3 中的所有条件。

为了完成最优拍卖的构造，我们设x为式（6.40）中的取值：

$$x_i(v)=P_i(v)\left[v_i+\sum_{\substack{j\in N\\ j\ne i}}e_j(v_j)\right]-\int_{a_i}^{v_i}P_i(v_{-i},s_i)ds_i$$

这一公式可以按如下方式重新以更为直观的方式写出来。对于除i之外的所有买主的任何估价组合v_{-i}，设

$$z_i(v_{-i})=\inf\{s_i\mid c_i(s_i)\ge v_0\text{且}c_i(s_i)\ge c_j(v_j),\forall j\ne i\} \tag{6.47}$$

则$z_l(v_{-i})$是i在给定v_{-i}下所有可赢得物品的买价中的下确界，故

$$P_i(v_{-i},s_i)=\begin{cases}1 & \text{若}s_i\ge z_i(v_{-i})\\ 0 & \text{若}s_i<z_i(v_{-i})\end{cases} \tag{6.48}$$

这给出

$$\int_{a_i}^{v_i}P_i(v_{-i},s_i)ds_i=\begin{cases}v_i-z_i(v_{-i}) & \text{若}v_i\ge z_i(v_{-i})\\ 0 & \text{若}v_i<z_i(v_{-i})\end{cases} \tag{6.49}$$

最后，式（6.40）变成

$$x_i(v)=\begin{cases}z_i(v_{\text{-}i})+\sum_{\substack{j\in N\\ j\ne i}}e_j(v_j) & \text{若}P_i(v)=1\\ \\ 0 & \text{若}P(v)=0\end{cases} \tag{6.50}$$

即买主i仅在他获得物品时才会付钱，且若他提交其可能最低的赢得价时，他支付的$\tilde{v}_i(v_{-i},z_i(v_{-i}))$是物品对于他的价值。

若所有的修正效应函数恒为零（即$e_i(v_i)=0$）且若所有买主是对称的$(a_i=a_j,b_i=b_j,f_i=f_j)$和正规的，则

$$z_i(v_{-i})=\max\{c_i^{-1}(v_0),\max_{j\ne i}v_j\} \tag{6.51}$$

即我们的最优拍卖变成了一个修正的 Vickrey 拍卖法（Vickrey，1961），其中卖主自己提交一个等于$c_i^{-1}(v_0)$的买价（注意：在这种对称情形，有$c_i=c_j$，且正规性保证有c_i是可逆的）且将物品交给最高出价买主，价格是第二高的报价。然而，仅当买主是对称的时且函数$c_i(\cdot)$是严格递增时，这一结论才成立。

譬如，假设对每个i和每个在 0 和 100 之间的v_i，$v_0=0$，每个$a_i=0$，$b_i=100$，$e_i(v_i)=0$且$f_i(v_i)=\frac{1}{100}$，则直接的计算给出$c_i(v_i)=2v_i-100$，它是v_i的增函数。所以卖主将把物品卖给最高出价的买主，价格为第二高出价，除非他自己提交一个$c_i^{-1}(v_0)=0+\frac{100}{2}=50$的买价。通过宣布一个保留价格 50，卖主冒着以概率$\left(\frac{1}{2}\right)^n$不能将物品售出即使某些买主愿支付高于$v_0$的价格买物品的风险（至少有某个$c^{-1}{}_i(v_i)\ge c^{-1}{}_i(v_0=0)=50,$ 物品才能被卖出，但这就要求$v_i\ge\frac{100}{2}=50,$ 由于v_i在$[0,100]$中均匀分布，故概率只有 0.5，物品不能被卖出的概率为0.5^n），但卖主仍然提高了其期望收益，因为他在物品售出时可以获得一个较高的价格。

于是，最优拍卖不是“事后”有效的。为了更加清晰地说明这一点，考虑上面的例子，并令其中的$n=1$。则卖主有估价$v_0=0$，且唯一的买主有一个取自[0,100]上均匀分布的估价。事后有效要求买主总能得到物品，只要其估价是正的。

但是，由于任何正的买价都会赢得物品。买主不会报出任何大于无穷小的买价。故卖主若不打算保留物品，则他只会得到零期望收益。事实上，卖主的最优政策是对于任何低于 50 的价格都拒绝出售物品，这会带给他 25 的期望收益。

更为一般地，当买主是不对称的时，最优拍卖有时甚至可能将物品卖给估价不是最高的一位买主。譬如，当$e_i(v_i)=0$且$f_i(v_i)=\frac{1}{(b_i-a_i)}$对于所有$[a_i,b_i]$中的$v_i$成立（没有修正效应时的一般性均匀分布情形），有

$$c_i(v_i)=2v_i-b_i$$

（$F(v_i)=\int_{a_i}^{v_i}f_i(v_i)dv_i=f_i(v_i)(v_i-a_i)=\frac{v_i-a_i}{(b_i-a_i)}$,由式（6.45)得到此式）

它是v_i的增函数。故在最优拍卖中，具有最高$c_i(v_i)$的买主将得到物品。若$b_i<b_j$，则即使$v_i<v_j$买主i也会赢得物品，只要$2v_i-b_i>2v_j-b_j$。事实上，对于估价上界较高的买主，最优拍卖是歧视他们的。这种歧视对于那些低于代表性估价的买主是不利的，而对于那些接近于较高b_j上界的是有利的。

6.5.5 一般情形下的最优拍卖

如果没有正规性假设，前述拍卖机制可能不是可行的，因为它将违反式（6.34）。为了将我们得到的解扩充至一般情形，我们需要一些更为仔细选择的定义。

买主i的分布函数$F_i:[a_i,b_i]\to[0,1]$是连续的和严格递增的，因为我们假定密度函数f_i总是严格正的。故$F_i(\cdot)$有反函数$F_i^{-1}:[0,1]\to[a_i,b_i]$，它也是连续的且严格递增。

对于每个买主i，我们现在定义四个函数，它们定义在单位区间[0,1]上。首先，

对于$[0,1]$中的任何q，设

$$h_i(q) = F_i^{-1}(q) - e_i[F_i^{-1}(q)] - \frac{1-q}{f_i[F_i^{-1}(q)]} \quad (6.52)$$
$$= c_i[F_i^{-1}(q)]$$

且设

$$H_i(q) = \int_0^q h_i(r)dr \quad (6.53)$$

再设$G_i:[0,1]\to R$为函数$H_i(\cdot)$的凸包；在 Rockafellar 的记号中（Ro-ckafellar, R. T., 1970，P36），有

$$G_i(q) = conrH_i(q)$$
$$= \min\{wH_i(r_1) + (1-w)H_i(r_2) \mid \{w, r_1, r_2\} \subset [0,1]\text{且} wr_1 + (1-w)r_2 = q\}$$

（6.54）

即$G_i(\cdot)$是$[0,1]$上的最高凸函数满足$G_i(q) < H_i(q)$对每个q成立。作为凸函数，G_i在除可数个点外是连续可微的且其导数为单调增的，定义$g_i:[0,1]\to R$满足

$$g_i(q) = G_i^{'}(q) \quad (6.55)$$

当该导数一旦被定义后，我们就可运用右连续方法将$g_i(\cdot)$扩充至$[0,1]$上所有点上。定义$\overline{c}_i:[a_i,b_i]\to R$满足

$$\overline{c}_i(v_i) = g_i(F_i(v_i)) \quad (6.56)$$

（可直接验证的是，在$c_i(\cdot)$是递增的正规情形，我们得到$G_i = H_i$，$g_i = h_i$，且$\overline{c}_i = c_i$）。最后，对所有估价向量v，设$M(v)$为其$\overline{c}_i(v_i)$是所有买主中最大的且大于v_0的那些买主的集合。

$$M(v) = \left\{i \mid v_0 \le \overline{c}_i(v_i) = \max_{j\in N} \overline{c}_j(v_j)\right\} \quad (6.57)$$

我们现在能陈述主要的结果：在最优拍卖中，物品总是被卖给有最高$\overline{c}_i(v_i)$的买主，只要$\overline{c}_i(v_i)$不低于v_0。

所以，我们可将$\overline{c}_i(v_i)$视为当买主i的出价为v_i时买主在卖主的最优拍卖中的先验水平（Priority level）。

定理 6.2　设$\overline{P}:V\to R^n$和$\overline{x}:V\to R^n$满足

$$\overline{P}_i(v) = \begin{cases} \dfrac{1}{|M(v)|} & \text{若} i \in M(v) \\ 0 & \text{若} i \notin M(v) \end{cases} \quad (6.58)$$

且

$$\overline{x}_i(v) = \overline{P}_i(v)\tilde{v}_i(v) - \int_{a_i}^{v_i} \overline{P}_i(v_{-i}, s_i)ds_i \quad (6.59)$$

对于V中所有v和N中所有i成立。则$(\overline{P},\overline{x})$是一种最优拍卖机制。

证明：

首先，利用分部积分，有下列方程

$$
\begin{aligned}
&\int_v [h_i(F_i(v_i)] - g_i[(F_i(v_i))]P_i(v)f(v)dv \\
&= \int_{a_i}^{b_i} [h_i(F_i(v_i)) - g_i(F_i(v_i))]\theta_i(P,v_i)f_i(v_i)dv_i \\
&= [(H_i(F_i(v_i)) - G_i(F_i(v_i)))\theta_i(P,v_i)]\Big|_{v_i=a_i}^{b_i} \\
&\quad - \int_{v_i=a_i}^{b_i} [H_i(F_i(v_i)) - G_i(F_i(v_i))]d\theta_i(P,v_i)
\end{aligned}
\tag{6.60}
$$

但 G_i 是 H_i 在 $[0,1]$ 上的凸包且 H_i 是连续的，故 $G_i(0)=H_i(0)$ 且 $G_i(1)=H_i(1)$。所以，上述最后一个表达式的内部项为零。

现在，回忆引理 3 中的最大化项式（6.39），利用式（6.60）有

$$
\begin{aligned}
&\int_v \sum_{i\in N}\left[v_i - e_i(v_i) - \frac{1-F_i(v_i)}{f_i(v_i)} - v_0\right]P_i(v)f(v)dv \\
&= \int_v \sum_{i\in N}\left[h_i(F_i(v_i) - v_0\right]P_i(v)f(v)dv \\
&= \int_v \sum_{i\in N}\left[\bar{c}_i(v_i) - v_0\right]P_i(v)f(v)dv + \sum_{i\in N}\int_v \left[h_i(F_i(v_i)) - g_i(F_i(v_i))\right]P_i(v)f(v)dv \\
&= \int_v \left[\sum_{i\in N}(\bar{c}_i(v_i) - v_0)P_i(v)\right]f(v)dv - \sum_{i\in N}\int_{v_i=a_i}^{b_i}\left[h_i(F_i(v_i)) - G_i(F_i(v_i))\right]d\theta_i(P,v_i)
\end{aligned}
\tag{6.61}
$$

现在，考虑定理中定义的那种 $(\bar{P},\bar{x})$。不难观察到，$\bar{P}$ 总是将所有的概率放在那些 $(\bar{c}_i(v_i)-v_0)$ 是非负和最大化的买主上。所以，对满足式（6.30）的所有 P 都有

$$
\int_v \left[\sum_{i\in N}(\bar{c}_i(v_i)-v_0)\bar{P}_i(v)\right]f(v)dv \ge \int_v \left[\sum_{i\in N}(\bar{c}_i(v_i)-v_0)P_i(v)\right]f(v)dv \tag{6.62}
$$

当然，$\bar{P}$ 本身并不满足概率条件（6.30）。

对于满足式（6.34）的所有 P（即对于 $\theta_i(P,v_i)$）是 v_i 递增函数，都有

$$
\int_{v_i=a_i}^{b_i}\left[H_i(F_i(v_i)) - G_i(F_i(v_i))\right]d\theta_i(P,v_i) \ge 0 \tag{6.63}
$$

因为 $H_i \ge G_i$。

为了看出 $\bar{P}$ 满足式（6.34），观察到 $\bar{c}_i(v_i)$ 是 v_i 的增函数，因为 F_i 和 g_i 都是增函数。故 $\bar{P}_i(v)$ 对于任何给定的 v_{-i} 来说是 v_i 的增函数，并且 $\theta_i(\bar{P},v_i)$ 也是 v_i 的增函数。所以 $\bar{P}$ 满足式（6.34）。

由于 G 是 H 的凸包，我们知道 $G<H$ 是水平的；即若 $G_i(r)<H_i(r)$，则 $\theta_i(\bar{P},v_i)$ 在 v_i 的某个领域为常数。这就意味着

$$
\int_{v_i=a_i}^{b_i}\left[H_i(F_i(v_i)) - G_i(F_i(v_i))\right]d\theta_i(\bar{P},v_i) = 0 \tag{6.64}
$$

将式（6.62）、（6.63）和（6.64）回代进式（6.61），我们就看到 $\bar{P}$ 在约束（6.34）和（6.30）下最大化式（6.39）。这一事实联同引理 3 就证明了定理。

为了对这些重要的 $\bar{c}_i$ 函数加以某种实践性的理解，考虑 $n=1$ 的特别情形；即假设仅有一个买主。则我们的最优拍卖问题就变为：

$$\overline{P}_i(v_1)=\begin{cases}1 & 若\overline{c}_1(v_1)\geq v_0\\ 0 & 若\overline{c}_1(v_i)<v_0\end{cases}$$

$$\overline{x}_1(v_1)=\overline{P}_1(v_1)\cdot\min\{s_1\mid\overline{c}_1(s_1)\geq v_0\}$$

即卖者将以如下价格出售物品

$$\overline{c}_1^{-1}(v_0)=\min\{s_1\mid\overline{c}_1(s_1)\geq v_0\}$$

且若买者不打算支付这一价格，他就将物品保留在手中。

因此，如果买者 i 是唯一的买主，则卖主将出售物品与 i 当且仅当 $\overline{c}_i(v_i)$ 大于或等于 v。换句话说，$\overline{c}_i(v_i)$ 是卖主个人估价 v_0 的最高水平，以致于如果所有买主都被迁移走，卖主将会把物品以 v_i 或比它更低的价格售给 i。

6.5.6 独立性假设

在前面，我们假定买主的估价之间是相互独立的随机变量。独立性假设是过强的假定，故我们现在考虑一个例子，并表明其中的最优行动看起来好象并非估价是互相独立的。

为简单计，我们考虑一个离散的例子。假设有两个买主，其中每一个的估价或者为 $v_i=10$ 或者为 $v_i=100$。假设 (v_1,v_2) 的联合概率分布为：

$$prob(10,10)=prob(100,100)=\frac{1}{3}$$

$$prob(10,100)=prob(100,10)=\frac{1}{6}$$

显然，这两个估价值是相互不独立的。若两个买主都有较高的估价 $(v_1=v_2=100)$，则以价格 100 将物品卖给其中某个买主，并以相等的概率随机地决定哪一个买主得到物品。若一个买主有较高的估价(100)和另一个有较低的估价(10)，则将物品卖给出价较高的(100)那位，且向报较低价格的买主索取 30（但他得不到物品）。若两个买者都只有较低的估价(10)，则将 15 单位货币给其中一位，且将 5 单位货币和物品交给另一位，同时随机地选择物品的获得者。

这个拍卖机制的产出函数 (P,x) 是：

$$P(100,100)=\left\{\frac{1}{2},\frac{1}{2}\right\}=P(10,10)$$

$$P(10,100)=(0,1),P(100,10)=(1,0)$$

$$x(100,100)=(50,50),x(10,10)=(-10,-10)$$

$$x(10,100)=(30,100),x(100,10)=(100,30)$$

这看起来是一个非常令人感到奇怪的拍卖，但事实上它是最优的。不难验证：在这个拍卖博弈中，说实话是一个纳什均衡，给定另一位买主说实话，任何买主都没有谎报估价的动机。进一步，物品总是转移给了估价最高者，并且无论其估价是高还是低，每一个买主从这个拍卖机制中获得的期望支付为零。故这个拍卖机制是可行的并

且它使得卖主将买主的所有消费者剩余都拿走。因此，这是一个最优拍卖机制，并且它带给卖主的期望收益为

$$U_0(P,x)=\frac{1}{3}(100)+\frac{1}{6}(130)+\frac{1}{6}(130)+\frac{1}{3}(-20)=70$$

为了看出为何这个拍卖机制如此运行，观察一下卖者实际上干着两件事情。首先，他将物品以最高的买主估价卖给估价最高的买主之一。其次，若买主说他的估价等于10，则这个买主就被强制性要求如下行事："如果其他买主的估价为100，则支付10；如果其他买主的估价为10，则获得150"。对于真实估价为10的买主，这使得买主的期望值为0，因为在另一位的估价为100时条件概率为$\frac{1}{3}$，另一位估价为10时条件概率为$\frac{2}{3}$。但是，当一位买主的估价为100但他却说谎并报出10的价格，则这使得他的期望值为$\frac{2}{3}(-30)+\frac{1}{3}(10)=-\frac{50}{3}$（因为现在他将对于其竞争对手的估价为100和10的两种情况分别赋予条件概率$\frac{2}{3}$和$\frac{1}{3}$）。这个对于说谎买主的负数期望值恰好抵消了买主以较低价格买到物品的说谎诱惑。

在独立性假设下，这个方法是不可能发挥的，因为每个买主关于其他买主的估价的条件概率为常数。但是，在一般的非独立情形，我们可以预测这个方法很可能出现。即卖主总可以通过以最高报价者的估价将物品卖给最高出价者将物品的价值全部获取，然后设置一个这种方法，若说实话则获得零期望值但若说谎则获负的期望值。当这种方法被仔细设计时，他们就能抵消为了以低价购得物品而说谎的冲动。

当然，在这个分析中，我们大量使用了风险中性假设。对风险中性买主，最优拍卖不太极端。在我们这个例子中所提出的拍卖博弈有不太好的第二个均衡，其中两个买者总是声称是低类型，尽管可设计出其他的最优拍卖机制，其中说实话均衡是唯一的（例如，将x变为：

$$x(100,100)=(100,100),x(10,10)=(-15,-15)$$

$$x(10,100)=(40,0),x(100,10)=(0,40)$$

而P保持上述的不变）。

是否在我们的例子中存在一种最优拍卖，它没有这样一种奇怪的性质即有时让卖者向买者支付价格？答案是否定的。如果我们附加一种约束即卖者绝不向买主支付货币（即所有的$x_i(v)\geq 0$），则没有一种可行的拍卖机制给卖者带来大于$66\frac{2}{3}$的期望支付。为了证明这一点，我们观察到如上例中那样估价值是有限的情形拍卖设计问题是一个线性规划问题。该问题的目标函数是$U_0(P,x)$，它对于P和x是线性的。可行约束有三种类型：概率约束$[P_i(v)\geq 0,\sum_i P_i(v)\leq 1]$，个人理性约束$[U_i(P,x,v_i)\geq 0]$和激励相容约束。所有这些约束是$P$和$x$的线性函数。故我们得到一个线性规划问题，且

对于我们的例子来说其最优值为 70，以及上述最优解。但若我们对所有 i 及 v 添加约束 $x_i(v) \geq 0$，则最优值跌落到 $66\frac{2}{3}$（对于这个例子来说）。为了获取这个“次优”值 $66\frac{2}{3}$ 和非负的 x，卖主应在 $v_1 = v_2 = 10$ 时保留物品，且否则卖主将物品卖与高出价买主从而获得 100。

6.6 混合战略的贝叶斯博弈解释

我们在第三章中对混合战略博弈的方法论基础展开了讨论，并指出其中蕴含的哲学问题。关于混合战略博弈，第三章中的讨论给出了多种可能的解释，其中主流的诠释来自 Harsanyi 的工作（Harsanyi, 1973）。这里，我们再重新回到这个问题上，并通过一个具体的例子说明 Harsanyi 解释的要点。在第三章中引出混合战略博弈时，我们假定信息是完全的，为了获得纳什均衡，我们扩充了局中人的战略空间，在纯战略空间基础上加入了随机选择纯战略的混合战略。这样，局中人在玩混合战略时，表现出来的特征是局中人选择纯战略上的不确定性。Harsanyi 的理解是：表面上看来似乎存在选择不确定性的混合战略，实际上是局中人在关于其他局中人类型把握上的信息不完全性的反映。当局中人的类型是私人信息时，其他局中人在纯战略选择上就存在不确定性。这是因为，如果对手的类型存在多种可能，特定局中人的最优纯战略也相应地有多种可能的选择。对于特定的局中人来说，倘若他不知道对手的类型，因而也就不知道对手的最优纯战略。因此，当他预测对手的纯战略选择时，他只能根据自己对对手类型的先验信息（对手类型的先验概率分布）作出对手即将选择的各种纯战略的预测，而这种预测是基于概率分布意义上的各种可能性。这样，对特定局中人来说，他好像感觉到对手是在随机地进行纯战略选择。给定这种预期，特定局中人就选择其期望支付最大化的纯战略，于是，当所有局中人都如此行为时，这构成了静态贝叶斯均衡。Harsanyi（1973）证明，当局中人类型的私人信息很小时，或博弈的信息不完全程度很弱时，这种本质上是静态贝叶斯博弈的博弈在数学形式上就等价于前面给出的完全信息混合博弈。当然，这种等价性是建立在数学家们所说的“几乎处处”的假设之上的，直观上说就是例外很少。下面，我们以第三章给出的“性别战”模型为例对此加以说明。在表 6.2 中，我们分别在男孩和女孩的支付函数上添加一个不确定变量 t_b 和 t_g，它们分别是男孩和女孩的私人信息[8]。

表 6.2　存在私人信息时的性别战博弈

		女：足球	女：芭蕾
男	足球	$2+t_b,1$	0,0
男	芭蕾	0,0	$1,2+t_g$

假设t_b和t_g是相互独立且服从$[0,x]$区间上均匀分布的贝叶斯随机变量。这里，两人的类型空间都为$[0,x]$，条件概率为$P(t_g \mid t_b)=P(t_b \mid t_g)=\frac{1}{x}$。

现在来看这一贝叶斯静态博弈。我们将证明，当男孩和女孩分别采用以下类型依存的战略时，将构成一个贝叶斯纳什均衡。

$$男孩：s_b=\begin{cases}足球 & t_b \ge c \\ 芭蕾 & t_b < c\end{cases} \tag{6.65}$$

$$女孩：s_g=\begin{cases}芭蕾 & t_g \ge P \\ 足球 & t_g < P\end{cases} \tag{6.66}$$

其中，c,P分别是$[0,x]$上的某个待定参数。

给定女孩的战略，男孩选择看足球的期望支付为

$$\frac{P}{x}(2+t_b)+\frac{x-P}{x}\cdot 0=\frac{P}{x}(2+t_b) \tag{6.67}$$

男孩选择看芭蕾的期望支付为

$$\frac{P}{x}\cdot 0+\frac{x-P}{x}\cdot 1=1-\frac{P}{x} \tag{6.68}$$

当且仅当$\frac{P}{x}(2+t_b)\ge 1-\frac{P}{x}$，男孩选择看足球是最优的，这等价于

$$t_b \ge \frac{x}{P}-3 \tag{6.69}$$

若令$c=\frac{x}{P}-3$，则男孩选择是最优的。

类似地，给定男孩的战略，女孩选择看芭蕾的期望支付为

$$\left[\frac{x-c}{x}\right]\cdot 0+\frac{c}{x}\cdot(2+t_g)=\frac{c}{x}(2+t_g) \tag{6.70}$$

女孩选择看足球的期望支付为

$$\left[\frac{x-c}{x}\right]\cdot 1+\frac{c}{x}\cdot 0=1-\frac{c}{x} \tag{6.71}$$

当且仅当

$\frac{c}{x}(2+t_g)\ge 1-\frac{c}{x}$时，女孩选择看芭蕾是最优的，这等价于

$$t_g \ge \frac{x}{c}-3 \tag{6.72}$$

若令$P=\frac{x}{c}-3$，则女孩选择是最优的，从而构成一个贝叶斯纳什均衡。联立这两个条件式

$$c=\frac{x}{P}-3$$

$P=\frac{x}{c}-3$

得一方程式：$P^2+3P-x=0$，解得：

$$P=\frac{-3+\sqrt{9+4x}}{2} \tag{6.73}$$

由对称性有 $P=c=\frac{-3+\sqrt{9+4x}}{2}$ （6.74）

其中的负根按模型含义舍去。

于是，男孩选择看足球的概率为

$$\frac{x-\frac{-3+\sqrt{9+4x}}{2}}{x}=1-\frac{-3+\sqrt{9+4x}}{2x}$$

当私人信息 x 很小时，即 $x\to 0$ 时，这一概率的极限为

$$\lim_{x\to 0}(1-\frac{-3+\sqrt{9+4x}}{2x})=1-\lim_{x\to 0}\frac{\frac{1}{2}\frac{4}{\sqrt{9+4x}}}{2}=\frac{2}{3}$$

这里使用了罗必塔法则。

由对称性，当 x 很小时女孩选择看芭蕾的概率极限也为 $\frac{2}{3}$，而这正是完全信息下性别战博弈的混合战略纳什均衡。

尽管这里的方法可以说是对混和战略博弈的贝叶斯博弈解释，但与前面对混和战略博弈的说明所不同的是，这里的男孩选择看足球的概率是指女孩对于男孩类型不确定性的认识。男孩自己是知道自己的类型的，因而不存在男孩自己随机地选择纯战略的问题。男孩子选择纯战略的随机性是女孩看来是如此的。但是我们在前面关于混和战略博弈的解释中，男孩是自己随机地选择纯战略的。当然，对于女孩来说，也存在类似的不同。

6.7 显示性原理

在上一节中，我们用到了由 R B.Myerson 提出并证明的著名的“显示性原理”。在这一节中，我们来给出这一原理由 Myerson 给出的原初表述及证明。“显示性原理”是博弈论中的一个著名的结果，也是 Myerson 作为一个博弈论专家在研究中所达到的一个引入注目的高度。这里，我们介绍 Myerson 在其原始论文（Myerson，1979）中给出的描述和证明，并采用他在这篇文章中使用的符号。

6.7.1 显示性问题的提出与表述

假定有一个仲裁者（arbitrator），他打算在一群行为主体（a group of individuals）

各自提出的要求的情况下完成一种群决策（collective choice），他对于这些主体的偏好和禀赋都只拥有不完全的信息。譬如，在拍卖活动中，拍卖人对于竞标者所愿支付的最高价格及他们的财富水平是不太清楚的，这个仲裁者不仅要努力摆平（Settling fairly）这些不同主体的可能是相互冲突的要求，而且他有时还会首先要搞清楚这些主体的真实偏好。当然，他可以要求其委托人告诉他自己他们需要知道什么，但若他不能迫使他们说实话的话，他就要面临某些主体可能说谎并试图由此操纵他的最终决策的问题。在这里，Myerson 提出了一种解决这个仲裁者问题的方法，其基础是由 Hurwicz 提出的激励相容（incentive—compatibility）概念（Hurwicz, 1972）和纳什提出的讨价还价解（Bargaining Solution）概念（Nash, 1950）。

在形式上，我们可以将仲裁者问题表述成如下的贝叶斯群决策问题（Bayesian Collective Choice Problem），其表达为

$$(C, A_1, A_2, \cdots, A_r, U_1, U_2, \cdots, U_n, P) \qquad (6.75)$$

其中，博弈的行为主体即除仲裁者之外的局中人为$1,2,\cdots,n$。C是仲裁者可选择的战略的集合。对于任一局中人i，A_i是局中人i的类型空间，即每一个$a_i \in A_i$都给出了局中人i相关特征的一个完整的描写，这些相关特征包括他的偏好，信念（belief）、能力及禀赋等等。U_i是局中人i的支付函数或效用函数，它将$C \times A_1 \times \cdots \times A_n$中的元素映射到实数域中，使得$U_i(c, a_1, a_2, \cdots, a_n)$是当$c \in C$被选定且$(a_1, a_2, \cdots, a_n)$为局中人的真实类型向量时局中人$i$将获得的支付。因此，假定$U_i$的取值是按 von Neumann—Morgenstern 效用函数方式计测的。最后，P是$A_1 \times \cdots \times A_n$上的一个概率分布，满足如下性质：$P(a_1, \cdots, a_n)$是仲裁者认为$(a_1, \cdots, a_n)$是局中人真实类型向量的概率。

为避开数学上的复杂性，假设C和所有的A_i都是非空有限集。

也就是说，每个局中人都只有有限种可能的类型，而对于这群主体来说也只有有限种群选择（Choice Options）。然而，我们不排除随机性选择的混合战略，即仲裁者可以选择C上的某些个概率分布，使得C上的实际选择按此分布随机地选出。

仲裁者关于这一群决策问题的解决方案的一种典型表达是：

仲裁者首先要求每一个局中人说出有关其自身类型的一些信息，然后在C中选出一个分配方案，或以一种概率分布在C中进行选择。在仲裁者的这一选择过程中，他使用了局中人告诉他的信息。为将这一过程模型化，我们定义一种选择机制（Choice mechanism），它是定义在$C \times (S_1 \times S_2 \times \cdots \times S_n)$上的一个实值函数$\pi$，对于反应集（response set）$S_1, S_2, \cdots S_n$的乘积$S_1 \times S_2 \cdots \times S_n$中的每一个向量$(s_1, \cdots, s_n)$和$C$中的所有$c$都有：

$$\sum_{c \in C} \pi(c \mid s_1, \cdots, s_n) = 1 \text{和} \pi(c \mid s_1, \cdots, s_n) \geq 0 \qquad (6.76)$$

这里，每个S_i是局中人i可能给予仲裁者问题的所有可能反应的集合，并且每一

个$\pi(c \mid s_1, \cdots, s_n)$是仲裁者在$(s_1, \cdots, s_n)$为他从主体群那里获得的反应组合时他将赋予选择$c$的概率。

如果仲裁者只是简单地向每一位局中人询问他的真实类型是什么，那么局中人i的反应集将为$S_i = A_i$，因为他的任何一个可能类型都将是一个回答。我们将称A_i为"标准反应集"（Standard response set）。除本小节和小节6.7.3之外，我们都将假定选择机制是建立在标准反应集上的。

假定每个局中人的反应是非合作地报告给仲裁者的。这样，当局中人i在其S_i中选择其反应s_i时，他并不知道其他局中人的反应是什么，并且他必须独立于其他局中人的决策过程完成他的选择。

有一些"真实反应"$T_i : A_i \to S_i$，使得$T_i(a_i)$是当局中人i的真实类型为a_i时局中人i的真实的反应（truthful response）。对于标准反应集，一个自然的真实反应就是$T_i(a_i) = a_i$。我们并不假定仲裁者有什么办法迫使局中人一定会给出真实的反映。完全知道其真实类型的局中人只有局中人自己，并且当局中人从说谎中可以谋得好处时没有人可以阻止他说谎。另一方面，如果说谎并无什么好处，我们假定局中人会说实话。

假定局中人总是接受仲裁者在C中所作出的最终推荐，即裁决是约束性的（binding）。

最后，假定由Harsanyi提出的"一致性条件"成立（Harsanyi, 1967）。为了在这个框架中表达这种一致性，需要以下定义。首先，每个局中人被给定了与仲裁者相同的信息（他知道基本结构(6.75)）以及额外的一些事实，即局中人i还知道自己的真实类型a_i。倘若仲裁者知道了局中人i的类型是a_i，则仲裁者将重新估计类型向量$\alpha = (\alpha_1, \cdots, \alpha_n)$的后验（Posterior）条件概率为：

$$P_i(\alpha_1, \cdots, \alpha_n \mid a_i) = \begin{cases} P(\alpha_1, \cdots, \alpha_n) / R_i(a_i) & 若\alpha_1 = a_i \\ 0 & 若\alpha_1 \neq a_i \end{cases} \tag{6.77}$$

其中$R_i(a_i)$是a_i的边缘概率（Marginal Probability），这是贝叶斯公式。

$$R_i(a_i) = \sum_{\substack{(\beta_1, \cdots, \beta_n) \in A_1 \times \cdots \times A_n \\ \beta_i = a_i}} P(\beta_1, \cdots, \beta_n) \tag{6.78}$$

一致性条件假定：$P_i(\alpha_1, \cdots, \alpha_n \mid a_i)$也是当局中人的类型为$a_i$时局中人$i$将对$(\alpha_1, \cdots, \alpha_n)$的概率所作出的估计。为何如此呢？一种辩护理由是：假设局中人的类型由某一众所周知的随机过程所决定，例如，如果局中人是从一个人群中随机挑选出来的，而这个人群中各种类型的比例是众所周知的[10]。

6.7.2 贝叶斯激励相容约束

在这一小节中，我们假设反应集都是标准反应集。也就是说，仲裁者向每个局中

人询问其类型是什么，局中人 i 的答案总是在 A_i 中。

因为仲裁者不能迫使局中人一定作出真实反应，他必须设计一种选择机制使得它不会激励说谎行为。给定一个选择机制 π，对于任意局中人 i 和任意的 $a_i \in A_i$ 和 $b_i \in A_i$，设

$$Z_i(\pi, b_i \mid a_i) = \sum_{\substack{\alpha \in A_1 \times \cdots \times A_n \\ c \in C}} P_i(\alpha \mid a_i)\pi(c \mid \alpha_{-i}, b_i)U_i(c, \alpha) \tag{6.79}$$

其中 $(\alpha_{-i}, b_i) = (\alpha_i, \cdots, \alpha_{i-1}, b_i, \alpha_{i+1}, \cdots, \alpha_i)$（回忆当 $\alpha_i \neq a_i$ 有 $P_i(\alpha \mid a_i) = 0$）。则 $Z_i(\pi, b_i \mid a_i)$ 是当局中人 i 预期所有其他局中人都说实话且他自己的类型为 a_i 且他在选择机制为 π 时向仲裁者报告说他的类型为 b_i 时，他将得到的条件——期望支付。

如果有

$$Z_i(\pi, a_i \mid a_i) \geq Z_i(\pi, b_i \mid a_i)\text{，任给 } i\text{，}a_i \in A_i\text{，}b_i \in A_i \tag{6.80}$$

则称建立在标准反应集基础上的选择机制 π 为“贝叶斯激励相容的”[11]。

如果一个选择机制是贝叶斯激励相容的，则没有一个局中人会期望当其他局中人都预计会说真话时他一个人说谎会获得好处。因此，当且仅当选择机制是贝叶斯激励相容的时，所有局中人都说真话才是一个均衡[12]。

因此，除非选择一个贝叶斯激励相容机制，通过标准反应集进行机制选择的仲裁者不能期望得到真实的反应。

如果选择了机制 π 及每个人都是说真话的，则局中人的条件期望支付为

$$v_i(\pi \mid a_i) = Z_i(\pi, a_i \mid a_i) \tag{6.81}$$

与机制 π 相关的条件期望支付配置（allocation of conditionally——expected payoff）为[13]

$$v(\pi) = [(v_1(\pi \mid a_1)_{a_1 \in A_1}, \cdots, (v_n(\pi \mid a_n)_{a_n \in A_n}] \tag{6.82}$$

向量 $v(\pi)$ 是当局中人知道仲裁者打算使用选择机制 π 及预期所有局中人都对仲裁者说真话时，局中人在其所有类型情况下将获得的期望支付。如果仲裁者能使用任意的选择机制及预期所有人都说真话，则我们可定义期望配置向量的“可行集”（feasible set）如下：

$$F = \{v(\pi) : \pi\text{是选择机制}\} \tag{6.83}$$

不幸的是我们知道，除非机制是贝叶斯激励相容的，否则不能指望反应一定是真实的。故我们必须将注意力限制在能达到贝叶斯激励相容机制的 F 的一个子集上去。于是，定义激励——可行的期望配置向量集为

$$F^* = \{v(\pi) : \pi\text{是贝叶斯激励相容的}\} \tag{6.84}$$

我们有如下定理：

定理 6.3　F^* 是 F 的一个非空凸紧子集。

证明：建立在标准反应集上的选择机制是定义在 $C \times A_1 \times \cdots \times A_n$ 上的一个实值函

数，且满足：

$$\sum_{c'\in C}\pi(c'|\alpha)=1 \text{ 和 } \pi(c|\alpha)\geq 0 \quad （6.85）$$

式（6.85）实际上在实变函数的理论框架中是按“几乎处处成立”来理解的，即不成立的点所构成的集合的勒贝格测度为零。假设有函数π_1和π_2都几乎处处满足条件式（6.85），λ是$[0,1]$中的数，则函数$\lambda\pi_1+(1-\lambda)\pi_2$也显然在$\pi_1$和$\pi_2$的满足式（6.85）的自变量集的交集上也满足式（6.85），π_1中不在此交集中的点必在π_2的一个零测集中，故这些点组成的集为一零测集。π_2中不在此交集中的点也必在π_1的一个零测集中，故这些点组成的集也为一零测集。所以，此交集外的所有点组成的集合也为零测集，因此，$\lambda\pi_1+(1-\lambda)\pi_2$也几乎处处满足式（6.85）。所以，建立在标准反应集上的所有选择机制构成定义在$C\times A_1\times\cdots\times A_n$上的所有实函数向量空间的一个凸子集。现在假设有一串机制$\{\pi_k\}$皆几乎处处满足式（6.85），且$\{\pi_k\}$几乎处处收敛于函数π。显然，π也几乎处处满足式（6.85），于是π也是一个建立在标准反应集上的选择机制。所以，建立在标准反应集上的所有选择机制构成定义在$C\times A_1\times\cdots\times A_n$上的所有实函数向量空间的一个紧子集。

据式（6.79），$Z_i(\pi,b_i|a_i)$是线性依赖于π的。不难看出，通过类似于上述分析的步骤，可证明满足式（6.80）的机制也构成一个紧凸子集，即贝叶斯激励相容机制的集合也构成一个紧凸集，因为它们几乎处处满足式（6.85）和（6.80）。激励相容机制的确是存在的。譬如，对所有$c\in C$和α，令$\pi(c|\alpha)=\frac{1}{|C|}$，则式（6.80）为恒等式，故式（6.80）成立，同时，条件式（6.84）也自然成立。所以，激励相容机制集是非空的，F和F^*也是非空的。

据式（6.81）、（6.82）知$v(\pi)$线性依赖于π。因为F^*是激励相容机制集的一个线性映射，故F^*必是配置空间中的一个非空紧凸集。F^*的非空性直接由激励相容机制的非空性得到。F^*的凸性是显然的，因为对于$\lambda\in[0,1]$和$\pi_1,\pi_2\in F^*$，则$\lambda v(\pi_1)+(1-\lambda)v(\pi_2)=v(\lambda\pi_1+(1-\lambda)\pi_2)$，而已知$\lambda\pi_1+(1-\lambda)\pi_2$是激励相容机制（已证激励相容机制的凸性），故$v(\lambda\pi_1+(1-\lambda)\pi_2)\in F^*$。$F^*$的紧致性直接由线性映射将紧集映为紧集的性质得到。

证毕！

众所周知的是：激励相容约束是一个有意义的约束，即F^*比F小得多（见Hurwicz(1972)和 Rosenthal(1979)），这一事实对于仲裁者有着重要的含义，因为他由此可以找到一种给所有局中人带来较高期望支付水平的选择机制。倘若某一机制π“严格”被占优于（Strictly dominated）另一机制π'，即

$$v_i(\pi|a_i)<v_i(\pi'|a_i) \quad \text{对所有 } i \text{ 和 } a_i\in A_i \text{ 成立} \quad （6.86）$$

则仲裁者必一采用机制π。

但是，对于许多群决策问题（见后面的例子），大多数非被占优的（undominated）或 F 的弱帕累托最优前沿可能在激励相容集 F^* 的外面。

因为我们不能在 F^* 的外面获得说真话均衡的配置，我们现在将提出这样一个问题：一个生成 F^* 外面的期望支付配置的选择机制能否具有一种均衡（允许预期有说谎行为）？在 6.7.3 小节中，我们将证明对此问题的答案是否定的，即对于任何选择机制来说，局中人的均衡反应行为总是带来一个 F^* 中的配置。

6.7.3 反应——计划均衡

在前述分析中，我们只考虑了建立在标准 (A_i) 反应集上的选择机制。我们现在转向考虑更加一般的情形。在整个这一小节中，我们仅假定每个局中人的反应集 S_i 是一个非空有限集。

局中人 i 的一个反应计划（response plan）是一个将每一类型 $a_i \in A_i$ 映射到其反应集 S_i 上的一个概率分布的一个函数 σ_i。也即是说，若 σ_i 是局中人 i 的反应计划，则 $\sigma_i(s_i \mid a_i)$ 是当其真实类型为 a_i 时局中人 i 将向仲裁者报告说他的类型是 s_i 的概率（因此，对所有 i 和 a_i，我们有 $\sigma_i(s_i \mid a_i) \geq 0$ 和 $\sum_{s_i^{'} \in S_i^{'}} \sigma_i(s_i^{'} \mid a_i) = 1$）。

若局中人 i 知道 a_i 是其真实类型及 $(\sigma_1, \cdots, \sigma_n)$ 列出了局中人在选择机制 π 下的反应计划，则局中人 i 的期望支付为

$$W_1(\pi, \sigma_1, \cdots, \sigma_n \mid a_i) = \sum_{\alpha \in A_1 \times \cdots \times A_n} \sum_{s \in S_1 \times \cdots \times S_n} \sum_{c \in C} P_i(\alpha \mid a_i) \cdot \left(\prod_{j=1}^{n} \sigma_j(s_j \mid a_j) \cdot \right) \pi(c \mid s) U_i(c, \alpha) \qquad (6.87)$$

由 $(\sigma_1, \cdots, \sigma_n)$ 生成的条件期望支付向量为[14]

$$W(\pi, \sigma_1, \cdots, \sigma_n) = \left((W_i(\pi, \sigma_1, \cdots, \sigma_n \mid a_i)_{a_i \in A_i})_{i=1}^{n} \right) \qquad (6.88)$$

注意，在这一小节中，我们不能将反应计划区分为“说真话”和“说谎”两大类，因为对于抽象的反应集 S_i，我们还未定义说真话的反应函数。

遵循 Harsanyi（1967-8），我们说对于选择机制 $\pi(\sigma_1, \cdots, \sigma_n)$ 是一个反应—计划均衡（response—plan equilibrium），如果对任何局中人 i 和类型 $a_i \in A_i$，对局中人 i 的每一个可能反应计划 $\sigma_i^{'}$ 有

$$W_i(\pi, \sigma_1, \cdots, \sigma_n \mid a_i) \geq W_i(\pi, \sigma_1, \cdots, \sigma_{i-1}, \sigma_i^{'}, \sigma_{i+1}, \cdots, \sigma_n \mid a_i) \qquad (6.89)$$

也即是说，一组反应计划构成一个均衡，如果没有一个局中人可以指望通过单方面的改变其计划而获利。

我们现在定义均衡——可行期望配置向量集为

$$F^{**} = \{ W_i(\pi, \sigma_1, \cdots, \sigma_n) : \pi\text{是选择机制且}(\sigma_1, \cdots, \sigma_n)\text{是对于}\pi\text{的反应}-\text{计划均衡} \} \qquad (6.90)$$

仲裁者可以通过承诺采用相应的选择机制 π 并且向局中人建议他们采用生成这一期

望支付配置的σ_i反应计划而带来任一F^{**}中的期望配置向量。可以确信局中人们会遵循他的建议，因为据均衡的定义，没有人会指望采用别的反应计划会带来更好的结果。

本小节的主要结果是均衡—可行性并不比激励相容性更为一般性。也就是说，对于任意选择机制的任何反应—计划均衡，存在一个等价的激励相容机制，它给所有类型的全体局中人带来相同的期望支付。因此，假定仲裁者将会选择建立在标准反应集上的激励相容机制不会失去一般性。

定理 6.4　$F^{**}=F^{*}$

证明：这一定理可以作为 Rosenthal 在 1979 年的文献中提出的定理 3 的一个推论而获得证明（Rosenthal，1979），但这里我们将直接采用 Myerson 的证明（Myerson，1979）。

如果$(\sigma_1,\cdots,\sigma_n)$是$S_1,\cdots,S_n$上的机制$\pi$的一个反应—计划均衡，则我们能定义一个$A_1,\cdots,A_n$上的一个等价的选择机制$\pi'$：

$$\pi'(c\mid\alpha)=\sum_{s\in S_1\times\cdots\times S_n}\pi(c\mid s)\cdot\left(\prod_{i=1}^{n}\sigma_i(s_i\mid\alpha_i)\right)$$

不难验证有

$$v(\pi')=w(\pi,\sigma_1,\cdots,\sigma_n)$$

所以，它们生成的配置是相同的。进而，π的均衡条件不等式（6.88）意味着π'的激励相容不等式（6.80）成立。因此，由$x=w(\pi,\sigma_1,\cdots,\sigma_n)\in F^{**}$就可推出$x=v(\pi')\in F^{*}$。故$F^{**}\subseteq F^{*}$。

为了验证还有$F^{*}\subseteq F^{**}$，只需观察到对于任意的激励相容机制π'，定义一个说真话的反应计划$\sigma_1^{'},\cdots,\sigma_n^{'}$如下：

$$\sigma_i^{'}(b_i\mid a_i)=\begin{cases}1 & 若 b_i=a_i\\ 0 & 若 b_i\neq a_i\end{cases}$$

它构成π'的一个反应计划均衡。故$x=v(\pi')\in F^{*}$意味着有$x=w(\pi',\sigma_1^{'}\cdots,\sigma_n^{'})\in F^{**}$。

证毕！

注释：

[1] 一般地有 $u_i = u_i(a_1,\cdots,a_i,\cdots a_n,\theta_i,\theta_{-i})$，但大多数场合 u_i 不包含 θ_{-i}，故有式（6.1）。

[2] 如果通过确定性效用函数构造不确定性下的效用函数，这一问题可以说是至今尚未完全解决。Von. Neuman-Morganstern 效用函数是通过将确定性效用函数的概率加权来获得不确定性下的效用函数，这一思路尽管有着直观意义上的依据，但给出的函数并不满足效用函数的性质，因为它在一般性的严格增函数复合下不能给出不变的偏好序，它只是在正线性变换下才给出不变的偏好序。支付函数应该在大多数场合就是效用函数，这里和第 3 章采用了 Von. Neuman-Morganstern 效用函数主要应该是基于其直观意义。

[3] θ_{-i} 从严格意义上看并非随机变量，但这里从局中人 i 的角度看它是在 H_i 上分布的随机变量。这种将主观上未知的变量作为随机变量处理的统计学背景被称为贝叶斯统计学，其方法论在学术界至今仍存在争议。但由贝叶斯（Bayes）发明的这种统计学在经济学中却得到大量应用。也就是从这个意义上，我们将这里介绍的博弈称为静态贝叶斯博弈。

[4] 不难验证，只有当我们做出这样的假设时，才可能获得线性最优战略。后面的结果将保证这一假设是成立的。

[5] 这种围标活动的组织和对“违规者”的惩罚有时甚至由黑社会来操纵，称为“黑道围标”。

[6] 这里的修正函数与 i 无关，意味着含有每一个局中人的修正函数是相同的假设。

[7] 为了说明 v_i 作为 i 的最初估价的合理性，我们可以假设修正函数具有零期望值，故

$$\int_{a_j}^{b_j} e_j(v_j) f_j(v_j) dv_j = 0$$

否则，其他局中人将修正对 v_j 的分布的信息。然而，对以下所要得到的结果来说，这一假设不是必需的。在没有这一假定的情形，仅仅是对 v_j 的理解有所不同。

[8] 注意：我们这里允许存在这样的可能性，即即使买者未得到物品，他也可能支付一定的价格，如购买门票等。

[9] 显然，这种假定的含义是当男孩非常喜爱足球时，即他从欣赏足球中获得的效用较大时，他会选择去看足球。对女孩来说也有类似的解释。

[10] 对于以下的分析来说，这一一致性假设实际上并不是必需的。事实上，我们只需应用条件概率函数 P_i 和边缘概率函数 R_i。反对一致性假设的读者可以将这些函数以如下方式理解：设 $P_i(\alpha_1,\cdots,\alpha_n \mid a_i)$ 是当局中人的真实类型为 a_i 时局中人 i 赋予

类型向量$(\alpha_1,\cdots,\alpha_n)$的主观概率（Subjective Probability），且设$R_i(a_i)$是仲裁者赋予局中人i的类型为a_i这一事实的先验（Prior）边际概率。在这种理解基础上，我们以下的所有结果都是成立的。

[11] D'Aspremont 和 Gerard—Varet（1977）曾建议引入形容词“贝叶斯”，以将此定义区别于由 Hurwicz（1972）给出的关于激励相容的另一较强的定义。这里，我们有时也省略掉形容词“贝叶斯”并同样指式（6.79），因为这里不会用到其它的有关“激励相容”的定义。

[12] 回忆我们曾假定局中人向仲裁者独立提交报告，因而没有一个局中人知道其他局中人向仲裁者说了什么。所以，每个局中人都是在给定其自身真实类型下比较其期望支付的大小来选择其最优反应。

[13] $v(\pi)$是$\sum_{i=1}^{n}|A_i|$个实数值的一个表列或向量，其标号是与各个A_i集合的不相交并集中的元素相 1-1 对应的。

[14] 这是一个有着$\sum_{i=1}^{n}|A_i|$个分量的向量，正如 6.7.2 中的$v(\pi)$向量那样。

第七章　不完全信息动态博弈

本章将在动态博弈中引入信息不对称因素，其博弈的内容被称为不完全信息动态博弈（uncomplete information dynamic game）。不完全信息动态博弈就其基本要素来看是前面引入的不完全信息概念与博弈的动态性质的一种综合。譬如，我们在处理不完全信息要素时是通过将某些局中人“类型”的不确定性作为信息不完全性的一种表征，这种方法将继续在本章中得以采用，即博弈中局中人面临的信息不完全性（无论它是指何种信息）将完全由某些局中人的“类型”的不确定性加以刻画。同时，作为动态博弈，正如我们在第五章中所指出的那样，“序贯理性”的思想将一直得到贯彻。我们将第五章中引入的“子博弈精炼均衡”的思想作类似的推广于不完全信息动态博弈。这种延续在逻辑上是必需的，因为一旦我们在不完全信息动态博弈中将信息不完全程度削减到零，则不完全信息动态博弈就自然应退化成一种完全信息动态博弈，其相应的精炼均衡概念就应回到子博弈精炼均衡。从这种意义上来看，不完全信息动态博弈的精炼均衡概念是子博弈精炼均衡概念的一种推广，正如不完全信息动态博弈应被视作完全信息动态博弈的一种推广一样。

7.1 精炼贝叶斯均衡

在本小节中，我们来构造不完全信息动态博弈的均衡概念，特别是贯彻了“序贯理性”原则的精炼均衡概念。

首先，博弈的纳什均衡是一种“僵持”状态的战略组合，当所有的局中人都选择该战略组合中给出的相应战略时，任何一个局中人都不会有单方面偏离这一选择的动机。作为动态博弈，一个战略是局中人在其可能进行行动选择的所有信息集上将作何选择的一整套规定或计划，而作为不完全信息博弈，这种规定或计划还是“类型依存”的，即不同类型的局中人将选择不同的战略规定。因此，一个不完全信息动态博弈的纳什均衡将是指这样的一种类型依存性的战略组合（或战略组合的族），当给定其他局中人的战略时（其他局中人的战略是类型依存的，所以，说给定其他局中人的战略即指给定其他局中人的战略与类型的依存关系），任一局中人在其任何类型下由该组合给出的类型依存战略给出的战略是其最优的。显然，这里还需要附加一个条件，即给定一局中人对其他局中人的类型分布的先验概率密度，否则他将无法对选择的“最优性”加以判断。这种概率分布或密度来自于博弈开始之前局中人所拥有的信息，故称为“先验”信息或“先验”概率。

于是，我们有如下定义：

定义 7.1　一个 n 人以不完全信息动态博弈的局中人类型空间 $H_1,\cdots,H_n$，条件概率 $P_i(\theta_{-i}\mid\theta_i), i=1,\cdots,n$，其中 θ_i 是局中人 i 的类型，$\theta_i\in H_i$，θ_i 的确定是通过 Harsanyi 转

换实现的，因而局中人 i 知道 θ_i 而其他除局中人 i 之外的局中人不一定知道 θ_i。一个类型依存的战略组合 $S^* = (S_1^*(\theta_1), \cdots, S_i^*(\theta_i), \cdots, S_n^*(\theta_n))$ 是一个纳什均衡，当且仅当有

$$S_i^*(\theta_i) \in \arg\max_{S_i(\theta_i)} \sum_{\theta_{-i}} P_i(\theta_{-i} | \theta_i) u_i(S_i(\theta_i), S_{-i}^*(\theta_{-i}), \theta_i, \theta_{-i}) \tag{7.1}$$

$i = 1, \cdots, n$

条件概率 $P_i(\theta_{-i} | \theta_i)$ 是先验的，因为它是博弈所给定的条件，来自博弈开始之前局中人 i 关于其他局中人类型的相关信息。当然，“自然”这个“局中人”并不包括在由下标 i 标记的 n 个局中人之中，但由 Harsanyi 转换所假定的局中人“自然”首先行动，它决定每一个局中人的类型，但除每个局中人自己能“观察”到自己的类型外，对于其他局中人的类型，他是只具有不完全信息的。按照式（7.1）定义的纳什均衡被称为贝叶斯纳什均衡，它在本质上与第 6 章定义的贝叶斯纳什均衡是一样的，因而并未体现出不完全信息静态博弈与不完全信息动态博弈的区别。正如我们在第五章和第三章所看到的，动态博弈与静态博弈的本质区别在于动态博弈均衡中存在对“序贯理性”的要求。这样，正如第五章中所做的那样，我们需要对定义 7.1 给出的纳什均衡加以精炼，以剔除含有不可置信承诺和威胁的均衡，这就是下面将要引入的“精炼贝叶斯纳什均衡”。

序贯理性在完全信息动态博弈中指的是局中人在任一子博弈上都选择最优的行动计划，而精炼均衡要求所有局中人的战略在任一子博弈上的限制都是其在给定其他局中人战略选择下的该子博弈上的最优战略，即纳什均衡战略。在不完全信息动态博弈中，信息集不一定是单结的，因而真子博弈可能不存在。此时，序贯理性指的是任一局中人在从其任一信息集开始的随后的博弈中所选择的行动计划都是最优的。对于任一局中人来说，当他处于某一信息集 h 上时，他对其他的每一个局中人的类型有一个概率判断，即他认为其他的每一个局中人的类型是某一特定类型的概率有一个判断；而给定其他每个局中人的一个特定类型情况下，他就知道其他每个局中人的战略是什么，即其他每个局中人在每个信息集上的行动选择是什么。此时，他也知道每个信息集的结构，且每个信息集都是单结的。这就是说，假定他认为其他每个局中人的类型是某个特定类型，博弈就变成完全信息的了。在完全信息动态博弈下，他当然知道此时他的最优行动计划是什么，即他知道他在此一单结信息集开始的子博弈上的最优行动计划是什么，它不过是使其支付最大化的行动计划而已。但事实上在不完全信息动态博弈中，他在此时并不准确知道其他局中人的类型是什么，但知道其他局中人的类型为每一种特定的类型组合的概率是多少。于是，假定所有局中人都是风险中性的，则他将根据这种概率分布来选择使他的期望支付最大化的行动计划。因为给定每一种其他局中人类型组合的概率分布及给定在其他局中人的每一种类型组合下他知道自己的最优行动计划，因而这种由概率密度加权的期望支付最大化行动计划是可以决定的。

设在局中人 i 的第 h 个信息集上[1]，局中人 i 自此信息集开始按行动顺序在随后的各

个信息集规定其行动选择构成一个自第 h 个信息集的一个行动计划，它是局中人 i 的战略在这个信息集及随后各个信息集上的限制，记为 $S_i(\theta_i)|_h$[2]，$S_i(\theta_i)$ 是局中人 i 的一个战略。给定其他局中人的战略组合 $S^*_{-i}(\theta_{-i})$，给定局中人 i 在第 h 个信息集上对其他局中人类型的概率密度 $\tilde{P}_{ih}(\theta_{-i})$，则类型为 θ_i 的第 i 个局中人的类型依存战略 $S^*_{-i}(\theta_i)$ 在其第 h 个信息集上的行动计划选择 $S^*{}_i(\theta_i)|_h$ 应满足如下条件：

$$S_i^*(\theta_i)|_h \in \arg\max_{S_i(\theta_i)|_h} \sum_{\theta_{-i}} \tilde{P}_{ih}(\theta_{-i}) u_i(S_i(\theta_i)|_h, S^*_{-i}(\theta_{-i}), \theta_i, \theta_{-i}) \qquad (7.2)$$

其中 $S_i(\theta_i)|_h$ 是局中人 i 在其第 h 个信息集上可能选择的行动计划（不是在第 h 个信息集上的行动选择），而 $S_i^*(\theta_i)|_h$ 指 $S_i^*(\theta_i) = S_i(\theta_i)|_h$ 的情形。

我们称 $\tilde{P}_{ih}(\theta_{-i})$ 为局中人 i 在其第 h 个信息集上的信念（belief）。

此时，我们称战略组合 $S^* = (s_1^*(\theta_1), \cdots, s_n^*(\theta_n))$ 是一个“精炼贝叶斯纳什均衡”。

显然，在精炼贝叶斯均衡中，局中人在不同信息集上的信息完全决定了其在不同信息集上的行动选择。从这种意义上看，精炼贝叶斯均衡实质上是局中人的信息之间的均衡。尽管如此，我们在表述一个精炼贝叶斯均衡时，通常是同时将信息与行动选择表达出来，即一个精炼贝叶斯均衡的表达中既要写出局中人在信息集上的行动选择，也同时标明他在信息集上的信念。

现在面临的问题是，如何给出局中人在信息集上的信息呢？

从原则上看，对信息的形成几乎是没有任何限制的，但是，按照贝叶斯统计学的假定，对于均衡路径上的信息集的信念形成可以通过贝叶斯法则加以约束，这就排除了均衡路径上信息的过多的任意性。所谓“均衡路径”，是指由那些发生的先验概率大于零的路径组成的信息集集合。当任一可能发生的局中人类型组合的先验概率大于零时，一旦这种类型组合成为真实的且成为局中人的共同知识，则它可能规定一个子博弈精炼均衡路径。当类型组合有多种可能性时，这些子博弈精炼均衡路径的集合将构成一个信息集集合，它就是“均衡路径”。

如果局中人 i 在其信息集 h 上观察到了其他局中人的行动组合 a^h_{-i}，则根据概率论中的贝叶斯公式：

$$\begin{aligned} prob(a^h_{-i}, \theta_{-i}) &= prob(a^h_{-i} | \theta_{-i}) prob(\theta_{-i}) \\ &= prob(\theta_{-i} | a^h_{-i}) prob(a^h_{-i}) \end{aligned} \qquad (7.3)$$

得到

$$\begin{aligned} prob(\theta_{-i} | a^h_{-i}) &= \frac{prob(a^h_{-i}, \theta_{-i})}{prob(a^h_{-i})} \\ &= \frac{prob(a^h_{-i}, \theta_{-i})}{\sum_{\theta_{-i}} prob(a^h_{-i} | \theta_{-i}) prob(\theta_{-i})} \end{aligned} \qquad (7.4)$$

式（7.4）中的 *prob* 表示概率。式（7.4）就是“贝叶斯公式”或“贝叶斯法则”（Bayes

Law），它将条件概率 $prob(\theta_{-i} \mid a_{-i}^h)$ 与先验概率 $prob(\theta_{-i})$ 联系起来。这里，在均衡路径上，条件概率 $prob(\theta_{-i} \mid a_{-i}^h)$ 就是前述信念 $\tilde{P}_{ih}(\theta_{-i})$。这是因为，在均衡路径上，$prob(a_{-i}^h) > 0$，但是，在非均衡路径上，$prob(a_{-i}^h) = 0$，贝叶斯公式的分母为零。由式（7.3），贝叶斯公式的分子也为零，因而贝叶斯公式在非均衡路径上给出的条件概率是 $\frac{0}{0}$ 型的数，是不确定的。所以，在非均衡路径，信念形成不受贝叶斯法则的制约，但也不是任意的，因为对于精炼贝叶斯均衡来说，非均衡路径上的信念与均衡路径上按贝叶斯法则决定的信念一起共同决定局中人在每一个信息集上的行动选择所构成的战略组合是精炼贝叶斯均衡。正是在均衡路径上我们按贝叶斯法则决定信念，所以称这种精炼均衡概念为贝叶斯纳什均衡。

在第六章中介绍的不完全信息静态博弈，实际上是经 Harsanyi 转换后将不完全信息静态博弈变成完全但不完美信息的动态博弈。这里，我们对不完全信息动态博弈的处理实际上也是通过 Harsanyi 转换将不完全信息动态博弈变成完全但不完美信息动态博弈。第六章实际上介绍的是一种被称为“二阶段”的完全但不完美信息动态博弈，第一阶段是“自然”选择局中人的类型，第二阶段是一个不对称信息的静态博弈，即局中人同时选择行动的博弈。

我们也称信念 $\tilde{P}_{ih}$ 为后验概率（Posterior probability）。

我们得出精炼贝叶斯均衡的正式定义如下：

定义 7.2 在不完全信息动态博弈 $G = \{u_1, \cdots, u_n, \tilde{P}_1, \cdots, \tilde{P}_n, H_1, \cdots, H_n\}$ 中，精炼贝叶斯纳什均衡是一个类型依存战略组合

$S^*(\theta_1, \cdots, \theta_n) = (S_1^*(\theta_1), \cdots, S_n^*(\theta_n))$ 及一个信念组合 $\tilde{P} = (\tilde{P}_1, \cdots, \tilde{P}_n)$，满足如下条件：

（1）$\tilde{P}$ 是先验概率 $P_i(\theta_{-i} \mid \theta_i)$ 的集合，即 $\tilde{P} = (P_1, \cdots, P_n)$，$P_i = P_i(\theta_{-i} \mid \theta_i)$，$\tilde{P}_i$ 是第 i 个局中人在其进行行动选择的信息集上所有信念组成的组合，记 $\tilde{P}_{ih}$ 为他在其第 h 个信息集上的信念；若局中人在信息集 h 上观察到的行动为 a_{-i}^h，则记 $\tilde{P}_{ih} = \tilde{P}_i(\theta_{-i} \mid a_{-i}^h)$，$i = 1, \cdots, n$。

（2）H_i 是局中人 i 的类型空间，$\theta_i \in H_i$ 是他的一个类型；$i = 1, \cdots, n$

（3）$u_i = u_i(s_1^*(\theta) \cdots s_n^*(\theta_n), \theta_i, \theta_{-i})$ 是局中人 i 的类型依存支付（效用）函数；$i = 1, \cdots, n$

（4）在第 i 个局中人的每一个信息集 h 上，有

$$S_i^*(\theta_i)\Big|_h \in \arg\max_{S_i(\theta_i)|_h} \sum_{\theta_{-i}} \tilde{P}_i(\theta_{-i} \mid a_{-i}^h) u_i(s_i(\theta)\Big|_h, s_{-i}^*(\theta_{-i}), \theta_i, \theta_{-i})$$

$i = 1, \cdots, n$

（5）在均衡路径上，$\tilde{P}_{ih}$ 是按照贝叶斯法则从先验概率 $P_i(\theta_{-i} \mid \theta_i)$，局中人 i 在信息集 h 上观察到的行动 a_{-i}^h 和 $S_{-i}^*(\theta_{-i})$ 导出的。

例 7.1 在图 7.1 中给出的一个不完全信息动态博弈中，“自然”N 首先选择局中人 1 的两种类型 $\{\theta_{11}, \theta_{12}\}$ 中的某一种，但局中人 2 对 N 的选择具有不完全信息，他只知道先验概率为 $P_2(\theta_{11} \mid \theta_2) = P_2(\theta_{12} \mid \theta_2) = 0.5$，局中人 2 的类型是对称的。

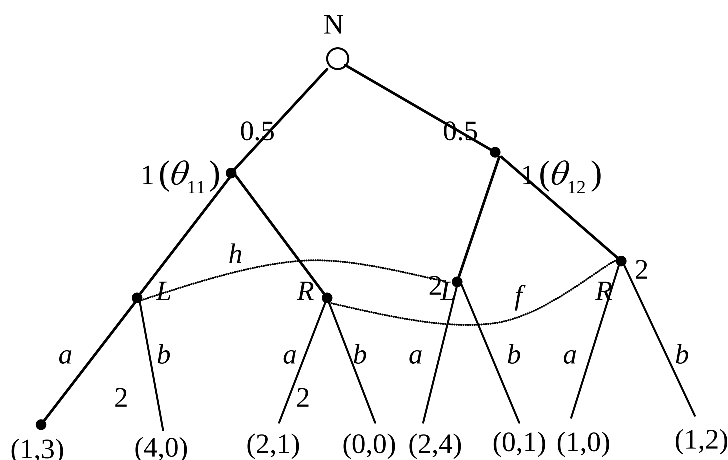


图 7.1　一个不完全信息动态博弈

如果局中人 1 的类型依存战略 $S_1^*(\theta_1)$ 为

$S_1^*(\theta_{11})=R$

$S_1^*(\theta_{12})=L$

试找出与此战略相对应的局中人 2 的一个类型依存战略 $S_2^*(\theta_2)$，使 $S^*=(s_1^*(\theta_1),s_2^*(\theta_2))$ 成为一个精炼贝叶斯均衡。

解：局中人 2 的类型是对称的。给定 $S_1^*(\theta_1)$，局中人 2 的 2 个信息集 h 和 f 都是均衡路径上的。根据贝叶斯法则，在信息集 h 上，局中人的后验概率为 $\widetilde{P}_2(\theta_{11}\mid L^h)=0$，$\widetilde{P}_2(\theta_{12}\mid L^h)=1$；同样，在信息集 f 上，有 $\widetilde{P}_2(\theta_{11}\mid R^f)=1$，$\widetilde{P}_2(\theta_{12}\mid R^f)=0$。给定这种信念，局中人 2 在 h 上的最优行动是 a，在 f 上的最优行动是 a。

给定局中人 2 在其信息集 h、f 上的上述信念及最优行动选择，类型为 θ_{11} 的局中人 1 将选择 R，类型为 θ_{12} 的局中人 1 将选 L，所以，$S_1^*(\theta_1)$ 和 $S_2^*(\theta_2)$ 构成一个精炼贝叶斯纳什均衡。我们用 $[(R,L),(a,a),(0,1),(1,0)]$ 表示这一均衡。其中，$S_1^*(\theta_1)=(R,L)$，$S_2^*(\theta_2)=(a,a)$，$\widetilde{P}_{2h}=(0,1)$，$\widetilde{P}_{2f}=(1,0)$。

在这个博弈中，还存在其它的精炼贝叶斯均衡。譬如，若 $S_1^*(\theta_{11})=S_1^*(\theta_{12})=L$。那么，给定局中人 1 的这一类型依存战略，局中人 2 在信息集 h 上的信息由贝叶斯法则有：

$\widetilde{P}_2(\theta_{11}\mid L^h)=\widetilde{P}_2(\theta_{12}\mid L^h)=0.5$

如果上述 $S_1^*(\theta_1)$ 是局中人 1 的精炼贝叶斯均衡战略，则局中人 2 的信息集 f 位于非均衡路径上，因而 $\widetilde{P}_{2f}$ 不受贝叶斯法则限制。但是，致使上述 $S_1^*(\theta_1)$ 成为局中人 1 的精炼贝叶斯均衡战略，就对 $\widetilde{P}_{2f}$ 的取值范围有另外的限制。

给定 $S_1^*(\theta_1)$，局中人 2 在 h 上选择 a 的期望支付为

$$0.5\times3+0.5\times4=3.5$$

而他选择 b 的期望支付为

$$0.5\times 2+0.5\times 1=1.5<3.5$$

因而他选择行动a。

此时，当局中人 1 在类型θ_{11}时选择 L 的支付为 1，在类型θ_{12}时选择 L 的支付为 2，欲使$S_1^*(\theta)$成为精炼贝叶斯均衡战略，就要求$\tilde{P}_2(\theta_{11}|R^f)$和$\tilde{P}_2(\theta_{12}|R^f)$取某些值使类型为$\theta_{11}$、$\theta_{12}$的局中人 1 都不会偏离上述选择。

局中人 2 在信息集f上选择行动a的期望支付为

$$\tilde{P}_2(\theta_{11}|R^f)\times 1+\tilde{P}_2(\theta_{12}|R^f)\times 0=\tilde{P}_2(\theta_{11}|R^f)$$

他选择行动b的期望支付为

$$\tilde{P}_2(\theta_{11}|R^f)\times 0+\tilde{P}_2(\theta_{12}|R^f)\times 2=2\tilde{P}_2(\theta_{12}|R^f)$$

当$\tilde{P}_2(\theta_{11}|R^f)\geq 2\tilde{P}_2(\theta_{12}|R^f)=2(1-\tilde{P}_2(\theta_{11}|R^f))$

局中人 2 在f上选a，此时有

$$\tilde{P}_2(\theta_{11}|R^f)\geq 2/3$$

否则，当$\tilde{P}_2(\theta_{11}|R^f)<2/3$时，局中人 2 在$f$上会选$b$，如果$\tilde{P}_2(\theta_{11}|R^f)\geq 2/3$，则类型为$\theta_{11}$的局中人 1 选$R$时的支付为$2>1$。这样，原有的$S_1^*(\theta)$就不是均衡战略了，因为类型为$\theta_{11}$的局中人 1 的最优行动是$R$而不是原有的$L$。

当$\tilde{P}_2(\theta_{11}|R^f)<2/3$，类型为$\theta_{11}$的局中人 1 选 R 时的支付为 0，故最优行动仍是 L。类型为θ_{12}的局中人 1 选 R 的支付为 1<2，最优行动仍为 L。所以，我们得到局中人 2 的最优战略为

$S_2^*(\theta)=(a,b)$，故精炼贝叶斯均衡为$S^*=[(L,L),(a,b),(0.5,0.5),\tilde{P}_2(\theta_{11}|R^f)<2/3,\tilde{P}_2(\theta_{12}|R^f)>1/3]$。

7.2 信号博弈：概念

在经济学的研究文献中，信号博弈作为一种特殊的不完全信息动态博弈得到了最为广泛的应用。正是信号博弈以一种十分特别的视角去理解很多令人感到迷惑的经济现象，信号博弈以及博弈论作为一种方法论才在主流经济学中产生了巨大影响。信号博弈通常描绘的是二个局中人之间的二阶段不完全信息动态博弈，其中，第一顺序行动的局中人的类型不为第二行动的局中人所知，他只知道第一顺序行动局中人的不同类型的先验分布概率。第二顺序局中人试图从他所观察到的第一顺序行动局中人所选择的行动中对其类型作出概率判断，从而选择自己的最优行动。在这种博弈中，后行动者主要关心的是先行动者的类型可能是什么，而先行动者也知道这一点。因而他有动机或者试图告诉后行动者他的真实类型，或者相反，他可能会试图欺骗后行动者，而努力将其有关他的类型的虚假信息告诉后行动者。当然，先行动者可以直接告诉后行动者他的类型是什么，但仅凭这种口头的承诺并不能使后行动者真正相信他所说的。如果他要后行动者相信他的话，他就必须作出一种努力，这种努力会使他蒙受一定的损失或存在一种成本。

这种成本是当他仅是这种类型时才能支付的，而如果他的类型不是这种类型，他不能承担这种成本。我们称这种成本支付是一种信号。通过它，先行动者能告诉后行动者他的真实类型。当然，说谎者也可以发出信号，并让后行动者难以准确判断其真实类型，如果这样做对先行动者是有利的话。譬如，文凭就是需要支付成本的一种信号，因为读书取得文凭是需要支付机会成本的一种活动，不同能力的人对这种成本的承受力是不同的。所以，雇主就可通过文凭去判断雇员的能力情况并据此支付不同的薪水。在金融市场上，如果一个企业需要在金融市场上融资，但投资者的对其真实的盈利能力具有不完全信息。于是，真正有高盈利能力的企业就可以通过向投资者支付较高的权益份额来将自己区别于低盈利能力的企业，从而让投资者识别出自己的真实类型而投资，而低盈利能力的企业由于对自己的真实盈利能力心知肚明，所以不敢模仿高盈利能力企业，它承诺的权益份额就较低，投资者不会将资金投入该企业。我们将在随后的内容中介绍这方面以及比这更多的内容，不过，在进行这种介绍之前，我们先应将信号博弈的概念形式化。

定义 7.3 在两个局中人 S（Sender，信号发送者）和 R（Receiver，信号接收者）的一个二阶段不完全信息博弈中。

（1）自然按照特定的先验概率密度 $P(\theta_i)$ 从 $H_1 = \{\theta_1, \cdots, \theta_n\}$ 中赋予发送者 S 特定的类型 θ_i，且有 $P(\theta_i) > 0$ 和

$$\sum_{i=1}^{n} P(\theta_i) = 1, \quad i = 1, \cdots, n$$

（2）发送者观察到 θ_i，随后从其类型依存的信号集 $M(\theta_i)$ 中选择一个信号 m_j 发送给接收者 R；

（3）接收者 R 观察到 m_j，然后从信号依存的行动集 $A(m_j)$ 中选择一个行动 a_R；

（4）发送者和接收者的支付为 $u_s(\theta_i, m_j, a_R)$ 和 $u_r(\theta_i, m_j, a_R)$。

在很多具体的信号博弈中，通常 M 与 θ_i 无关，A 与 m_j 无关。

按照精炼贝叶斯均衡的要求，在均衡情况下，上述过程中的（2）、（3）是按照发送者和接收者的最优行为进行的。

譬如，在接收者观察到信号 m_j 的信息集上，他的信念 $\tilde{P}(\theta_i \mid m_j)$ 满足

$$\sum_{i=1}^{n} \tilde{P}(\theta_i \mid m_j) = 1$$

他选择的最优行动 $a^*(m_j) \in A(m_j)$ 满足

$$a^*(m_j) \in \arg\max_{a \in A(m_j)} \sum \tilde{P}(\theta_i \mid m_j) u_r(\theta_i, m_j, a)$$

在发送者的单结信息集上，他的均衡信号选择满足

$$m_j^* \in \arg\max_{m_j \in M(\theta_i)} u_s(\theta_i, m_j, a^*(m_j))$$

我们可以将发送者的类型集H_1分为若干个子集h_j的不相交并集：

$H_1=\bigcup_j h_j$， $h_{j1}\bigcap h_{j2}=\phi$， $j_1\neq j_2$，ϕ为空集。

如果当$\theta_{j1},\theta_{j2}\in h_j$时有$m_{j1}^*=m_{j2}^*$，则称

该均衡为“准分离均衡”（Semi—Separating equilibrium）。如果对所有的类型θ_i，发送者发送相同的均衡信号，即$h_j=H_1$，则称为“混同均衡”（Pooling equilibrium）。当对每一个h_j，h_j只含有唯一一个元素时，也就是说，不同类型的发送者发送不同的信号时，称为“分离均衡”（Separating equilibrium）。当然，信号博弈不排除局中人随机地选择信号或行动，此时称这种混合战略为“杂合战略”（hybrid strategies）。

上述的例 7.1 就是一个信号博弈，我们来看看其纯战略精炼贝叶斯均衡有哪些。对于信号博弈，我们采纳 Gibbons（R.Gibbons, 1998）的方法，将博弈树画成更加方便于分析的一种形式。见图 7.2。显然，我们在例 7.1 中事实上已经找到了一个分离均衡$[(R,L),(a,a),(0,1),(1,0)]$ 和 一 个 混 同 均 衡 $[(L,L),(a,b),(0.5,0.5),\tilde{P}_2(\theta_{11}\mid R^f)<2/3,\tilde{P}_2(\theta_{12}\mid R^f)>1/3]$。

我们再来看看该博弈是否还存在其它的纯战略均衡。

显然，就纯战略均衡来看，只需探索其所有的分离和混同均衡即可。

除了例 7.1 已经给出的两个纯战略均衡外，有待确认的是可能的混同均衡(R,R)和分离均衡(L,R)，即类型分别为θ_{11}、θ_{12}的局中人 1 都同时选择行动 R 或类型为θ_{11}的局中人 1 选择行动 L，类型为θ_{12}的局中人 1 选择行动 L。

对于混同均衡(R,R)，如果它真是一个均衡，则在信息集h上，由于是非均衡路径，信念不受贝叶斯法则限制，但要使(R,R)是一个精炼均衡，对信念仍然存在限制。这是因为，在信息集f上，局中人 2 不能从观察到的行动 R 上获取任何关于局中人 1 类型的除先验概率之外的额外信息，故后验概率就等于先验概率，即

$\tilde{P}_2(\theta_{11}\mid R^f)=\tilde{P}_2(\theta_{12}\mid R^f)=0.5$

当局中人 2 选择行动a时，其期望支付为

$0.5\times1+0.5\times0=0.5$

而当他选择行动b时，期望支付为

$0.5\times0+0.5\times2=1>0.5$

所以，局中人 2 在信息集f上选择行动b。

现在，我们回到信息集h。因为当局中人 1 选择 R 时，他知道局中人 2 会选择行动b，故对于类型分别为θ_{11}、θ_{12}的局中人 1 来说，他们选择 R 时的支付分别为 0 和 1。如果(R,R)是精炼均衡，他们就不会选择 L，即一旦选择了 L，支付应分别不大于 0 和 1。显然，对于类型为θ_{11}的局中人 1 来说，一旦他选择了 L，无论局中人 2 作为选择，支付一定是大于 0 的（支付分别在局中人 L 选a或b时为 1 和 4）。所以，(R,R)不会是精炼

均衡。

再看分离均衡(L,R)是否是一个精炼均衡。此时，信息集h和f都在均衡路径上。在h上，局中人 2 会选择a，在f上，局中人 2 会选择b。给定这种可能，类型为θ_{12}的局中人 1 之最优选择为 L 而不是R。所以，(L,R)不会是一个精炼均衡。因而例 7.1 中的博弈只有两个纯战略均衡。它们分别是例 7.1 中已找出的一个混同均衡和一个分离均衡。

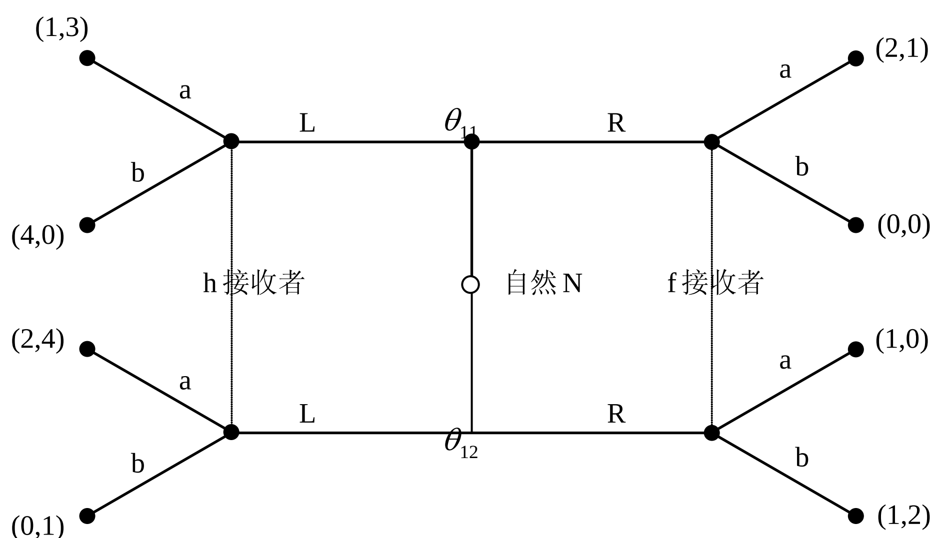


图 7.2　一个信号博弈

7.3 信号博弈：经济学理论中的几个应用模型

7.3.1 斯宾塞就业市场中的文凭信号模型

斯宾塞运用信号博弈模型对文凭的功用作了一种博弈论的解释，即文凭具有一种揭示雇员真实能力的信号传递功能（Spence, 1973）。

斯宾塞作为三位信息经济学的创始人之一而荣获 2001 年的诺贝尔经济学奖，[3]他所作出的这一模型事实上开创了广泛运用不完全信息动态博弈描述经济现象的先河。斯宾塞教授也是信号博弈研究的先驱者之一，他还是最早给出如精炼贝叶斯均衡等均衡概念定义的学者之一。斯宾塞模型是劳动经济学中的一个重要成果，它使我们对教育的功用有更多的认识。

模型假定

该模型描述的是一个信号博弈，其中有两个局中人，一个是雇员，其发出信号是教育水平或文凭，记为$e \in E$，我们用非负实数测度e，即$e \geq 0$。假设雇员的类型是其工作能力，分为高(H)和低(L)两种，并且他的类型是私人信息。设雇员类型为 H 的先验

概率为 P。另一个局中人是劳动力市场，它是由竞争性市场上的许多企业组成的。

设企业给予雇员的工资率为 w，θ 为雇员的类型（能力），$c(\theta,e)$ 为类型为 θ 的雇员在教育水平（学习努力程度）为 e 时的成本，$y(\theta,e)$ 为类型为 θ 且教育水平为 e 的的雇员的（边际）产出。当雇员被企业雇佣时，雇员的支付为 $w-c(\theta,e)$，而企业的（边际）支付为 $y(\theta,e)-w$。

一个基本的假设是：低能力的雇员与高能力雇员相比，要取得同样的教育水平（文凭）需花费较大的成本。我们用如下假设来刻画这种条件，即低能力雇员受教育的边际成本高于高能力雇员，即对所有的 e 有：

$$\frac{\partial c(L,e)}{\partial e} > \frac{\partial c(H,e)}{\partial e} \tag{7.5}$$

雇员的无差异曲线由如下方程刻画：

$$w-c(\theta,e)=R \tag{7.6}$$

其中 R 为雇员的常数净支付。当 $R=0$ 时，我们得到一条特定的无差异曲线。

$$w=c(\theta,e) \tag{7.7}$$

假定类型为 θ，教育水平为 e 的劳动市场是竞争性的，因而在信息完全下有式（7.7），它也是信息完全时类型为 θ 的文凭需求曲线。式（7.5）表明，低能力雇员的文凭需求曲线比高能力的文凭需求曲线陡一些，见图 7.3。

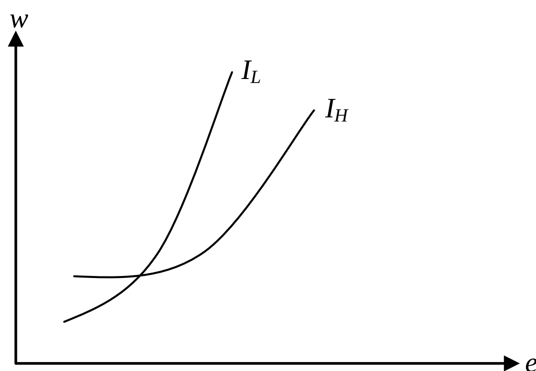


图 7.3　不同类型雇员的无差异曲线

图 7.3 中的无差异曲线既是雇员文凭的需求曲线，又是雇员对企业的文凭供给曲线。还假设企业之间在劳动市场上是竞争性的，因而在信息完全下边际利润为零，即

$$y(\theta,e)=w \tag{7.8}$$

在这个博弈中，文凭或教育水平是雇员向企业发出有关其类型的信号，w 是企业接收到这一信号后决定给予雇员的工资率。

当信息不完全时，如果企业观察到的信号为 e，则 $\widetilde{P}(H\mid e)$ 为企业认为雇员能力为 H 的后验概率，假定企业是风险中性的，则一个雇员的期望产出为

$$\widetilde{P}(H\mid e)y(H,e)+(1-\widetilde{P}(H\mid e))y(L,e) \tag{7.9}$$

由企业之间的竞争性，企业的行动选择是[4]

$$W(e)=\widetilde{P}(H\mid e)y(H,e)+[1-\widetilde{P}(H\mid e)y(L,e)] \qquad (7.10)$$

当信息完全时，$W(e)=y(\theta,e)$，于是能力为θ的雇员将选择满足如下条件的e：

$$\max_{e}[y(\theta,e)-c(\theta,e)] \qquad (7.11)$$

设式（7.11）的最优解为$e^{*}(\theta)$，在劳动市场和企业之间是竞争性的条件下，$y(\theta,e^{*}(\theta))=c(\theta,e^{*}(\theta))$，且$\dfrac{\partial y(\theta,e^{*})}{\partial e}=\dfrac{\partial c(\theta,e^{*})}{\partial e}$（雇员效用最大化条件，即式（7.11）），见图 7.4。

注意，图 7.4 中的生产函数$y(\theta,e)$是向上倾斜的，这说明在同样的能力水平下，获得较多教育将提高雇员的生产率。

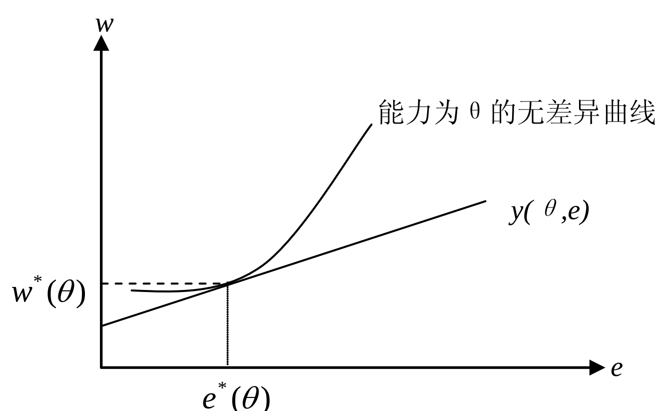


图 7.4　完全信息下的最优教育水平

现在，我们在不完全信息假定下展开博弈分析。

作为信号博弈，该模型在一定的条件下分别存在分离均衡、混同均衡等。我们先看看这两类均衡的情况。对于分离均衡来说，要求低能力雇员不能模仿高能力雇员，即低能力雇员如果模仿高能力雇员取得高学历文凭，则即使因此而获取高工资率$W^{*}(H)$也不能补偿其过高的成本，于是有

$$W^{*}(L)-c(L,e^{*}(L))>W^{*}(H)-c(L,e^{*}(H)) \qquad (7.12)$$

见图 7.5。

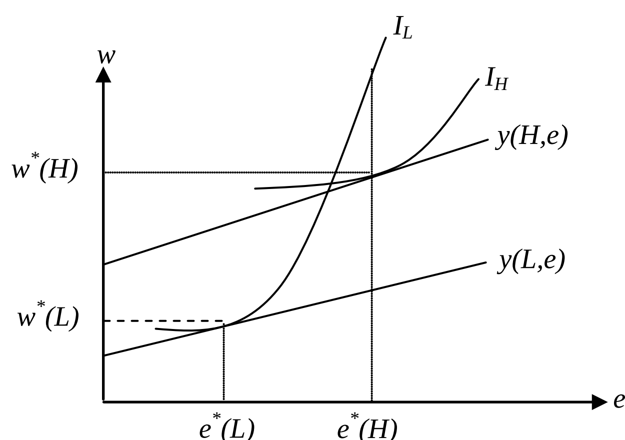


图 7.5　分离均衡条件

在混同均衡情形，低能力雇员模仿高能力雇员所耗成本小于高学历带来的收益。

$$W^*(L)-C[L,e^*(L)]<W^*(H)-C[L,e^*(H)] \quad (7.13)$$

见图 7.6。

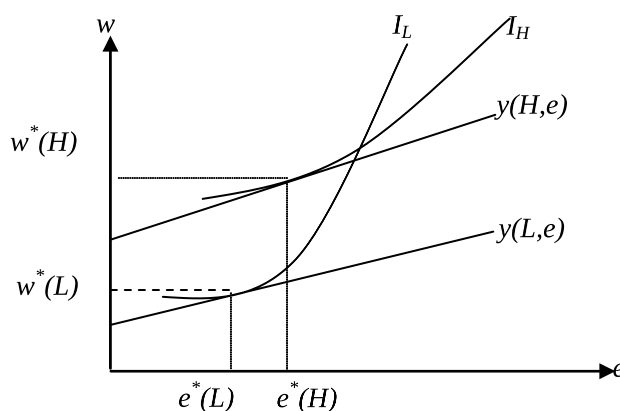


图 7.6　混同均衡条件

对于混同均衡，设两种类型的雇员都选择同样的教育水平 e_p，此时，企业在观察到 e_p 后有 $\tilde{P}(H|e_P)=P$， P 是雇员类型为 H 的先验概率。按照前述假设，此时企业的最优工资率选择为

$$W_P=Py(H,e_P)+(1-P)y(L,e_P) \quad (7.14)$$

在非均衡路径上，令 $\tilde{P}(H|e)=0$， $e\neq e_P$。于是有

$$\tilde{P}(H|e)=\begin{cases}P & e=e_p\\ 0 & e\neq e_P\end{cases} \quad (7.15)$$

由式（7.10）有企业的最优行动选择为

$$W(e)=\begin{cases}w_P & \text{在}e=e_P\text{的信息集上}\\ y(L,e) & \text{在}e\neq e_P\text{的信息集上}\end{cases} \quad (7.16)$$

给定企业的反应，能力为 θ 的雇员选择由以下条件决定的 e

$$\max_e[w(e)-c(\theta,e)] \quad (7.17)$$

当 $e_P=e^*(H)$ 时，在图 7.7 中，对于类型为 H 的雇员，当他选择 e_p 时，他所处的无差异曲线为 I_H，而当他选择 $e\neq e_P$ 时，他所处的无差异曲线为 $y(L,e)$ 上的点所处于的无差异曲线，显然效用小于前者，故选 e_p 为最优的。

对于类型为 L 的雇员，当他选择 e_p 时，处于无差异曲线 I_L，若他选择 $e\neq e_P$，则处于过 $y(L,e)$ 上的无差异曲线，显然选 e_p 是最优的，因为选 $e\neq e_P$ 中的最大化支付选择是 $e^*(L)$，过 $[e^*(L),w^*(L)]$ 的无差异曲线在过 $[e_P,w_P]$ 的无差异曲线的下方。显然，对于图 7.7 中的那种 $e_P=e^*(H)$，以及图中的那种无差异曲线和生产函数来说，雇员选择信号 $e_P=e^*(H)$ 构成一个混同均衡。从数学关系上看，混同均衡并不一定要求 $e_P=e^*(H)$，

还存在其它许多e_p是混同均衡信号。

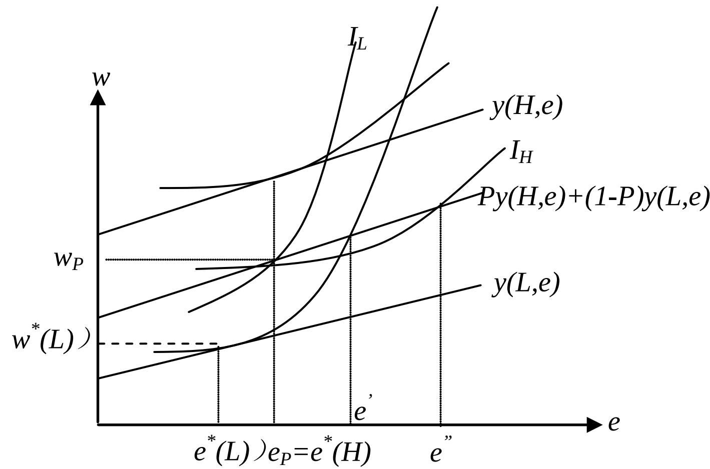


图 7.7 混同均衡

即使对于图 7.7 中那样的无差异曲线和生产函数，也还存在许多其它的混同精炼均衡，其中某些混同均衡之间可能选择的教育水平不同，但另一些混同均衡之间的差别仅在于非均衡路径上的信念不同。

显然，在图 7.7 中，只要选择e^*与e'之间的教育水平，也不难看出会构成另外的无限多个混同均衡。

我们还可看出，仅仅改变非均衡路径上的信息，也可以获得另外的混同均衡。

我们将式（7.15）修改为：

$$\tilde{P}(H\mid e)=\begin{cases}0 & e\le e''\text{但}e\ne e_P\\ P & e=e_P\\ P & e>e''\end{cases}\qquad（7.18）$$

其中e_P是上述位于$e^*(H)$与e'之间的任一信号。企业战略为：

$$W(e)=\begin{cases}y(L,e) & e\le e''\text{但}e\ne e_P\\ w_P & e=e_P\\ w_P & e>e''\end{cases}\qquad（7.19）$$

类型为 H 的雇员选e_P时，位于过(w_P,e_P)的无差异曲线I_H，但当他选择$e\le e''$但$e\ne e_P$时，位于过$y(L,e)$的无差异曲线，当他选择$e>e''$时，位于I_H下方的无差异曲线。故选e_P是最优的。对于类型为 L 的雇员，他选e_P时位于无差异曲线I_L上，但当他选$e\le e''$且$e\ne e_P$时，位于过$y(L,e)$的无差异曲线，当选$e>e''$时，位于I_L下方的无差异曲线上。故选e_P是最优的。

下面，我们来看看分离均衡情形，见图 7.5。显然，最自然的分离均衡中的雇员战

略为$[e(L)=e^*(L),e(H)=e^*(H)]$，此时，企业在观察到信号后的后验概率为

$$\tilde{P}(H\mid e^*(L))=0\text{和}\tilde{P}(H\mid e^*(H))=1 \qquad (7.20)$$

据式（7.10）有：

$$W(e^*(L))=W^*(L)\text{和}W(e^*(H))=W^*(H) \qquad (7.21)$$

在非均衡路径上的信念规定如下：

$$\tilde{P}(H\mid e)=\begin{cases}0 & e<e^*(H)\\ 1 & e\geq e^*(H)\end{cases} \qquad (7.22)$$

企业的最优选择为：

$$W(e)=\begin{cases}y(L,e) & e<e^*(H)\\ y(H,e) & e\geq e^*(H)\end{cases} \qquad (7.23)$$

对于类型为 H 的雇员，当他选$e^*(H)$时，位于无差异曲线I_H上；当选$e>e^*(H)$，位于无差异曲线I_H下方的无差异曲线上，故选e劣于$e^*(H)$。当选$e<e^*(H)$时，收入为$y(L,e)$，此时他位于无差异曲线I_H的下方，由图 7.5 知道他处于过$y(L,e)$曲线的较低位置的无差异曲线上。。已知他在$e>e^*(H)$时的收入为$y(H,e)$，其效用也不如选$e^*(H)$，故此时效用仍不如选$e^*(H)$。所以，他选$e^*(H)$是最优的。

对于类型为 L 的雇员，当他选$e<e^*(H)$，收入为$y(L,e)$，必小于选$e^*(L)$时的效用，因$e^*(L)$是他的工资函数为$w=y(L,e)$时的的最优努力水平。当他选$e\geq e^*(H)$时，收入为$y(H,e)$，净收益即支付为

$$y(H,e)-c(L,e)$$

据图 7.5，该净收益（为负）显然小于选$e^*(L)$的净收益（为零）。故选$e^*(L)$是最优的。

如果不假设图 7.5 那样的几何关系，而是在图 7.6 的情形下，我们来寻找分离均衡。此时，由于存在低能力雇员模仿高能力雇员的倾向，高能力雇员要阻止低能力雇员的模仿，从而形成分离均衡，就必须要求高能力雇员付出更高的代价，取得更高的教育水平，使低能力雇员难以模仿。企业观察到这种较高水平的教育成本，也因此而知道他是高能力雇员，并给予较高的工资率予以奖励。这就象雄性动物在求偶时为了向雌性动物表明自己是最合适的传宗接代对象（有最好的体质）而相互打斗一样，胜者需要付出额外的成本。中国武侠小说中常见到两个狭路相逢的好汉在一场恶战之前，通常要各自表演一番武功，如用手劈砸砖石等绝技。这是高手为避免由于双方信息不完全而导致一场双方都头破血流的打斗所发出的信号。譬如，当一个好汉见对方的绝技自己不能模仿时，可能就明白了对手是比自己技高一筹的好汉，他会选择弃战或甘拜下风，而避免更惨的结局。

由此，我们料想高能力雇员选$e_s>e^*(H)$才能构成分离均衡，见图 7.8。

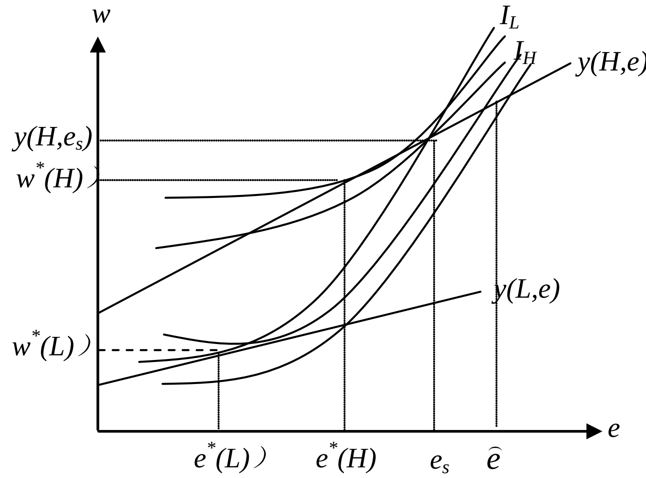


图 7.8 存在信号发送成本的分离均衡

在$e^*(H)$到e_s之间的教育水平e，如果低能力雇员效仿高能力雇员可令企业误认为他是高能力雇员，则低能力雇员有动机如此做。但是，当高能力雇员的信号等于e_s时，低能力雇员在模仿高能力雇员选e_s与暴露自己类型的选择$e^*(L)$之间是无差异的，可假设此时他选$e^*(L)$。下面给出信息集上的信念，并证明它与雇员的战略一起构成分离均衡。

雇员战略为

$$e(\theta)=\begin{cases}e^*(L) & \theta=L\\ e_s & \theta=H\end{cases} \tag{7.24}$$

当企业观察到信号后，后验概率为

$$\tilde{P}(H\mid e)=\begin{cases}0 & e<e_s\\ 1 & e\geq e_s\end{cases} \tag{7.25}$$

企业的最优选择为

$$W(e)=\begin{cases}y(L,e) & e<e_s\\ y(H,e) & e\geq e_s\end{cases} \tag{7.26}$$

对于低能力雇员，选择$e^*(L)$与选e_s的净收益都为零，而选其它的信号净收益为负，在前述假定下选$e^*(L)$为最优。对于高能力雇员，当$e\geq e_s$时，无差异曲线在I_H的下方；而当$e<e_s$时，无差异曲线显然也在I_H下方，故选e_s为最优。

该博弈还存在其它的分离均衡。在某些分离均衡中，高能力雇员选择不同的教育水平，而低能力雇员总选$e^*(L)$。另外一些分离均衡是高能力雇员总选e_s，低能力雇员总选$e^*(L)$，但非均衡路径上的信念不同。作为前者的一个例子，令$\hat{e}$为高于e_s的教育水平，但又不足以高到使高能力雇员不愿意选择$\hat{e}$，而宁愿被认为自己是低能力的，即图 7.8 中的$\hat{e}$。如果在图 7.8 中，用$\hat{e}$替换$\tilde{P}(H\mid e)$和$W(e)$表达式中的e_s，则由此形成的企业信

念和战略与雇员战略$[e(L)=e^*(L),e(H)=\hat{e}]$一起，构成一个分离均衡。作为后者的一个例子，令企业对严格处于$e^*(H)$和e_s之间的教育水平的信念严格大于 0，为一个足够小的正数，使得据此得出的战略$W(e)$严格处于低能力雇员通过点$[e^*(L),w^*(L)]$的无差异曲线的下方。

最后，我们来讨论杂合均衡。我们考察一种特定的杂合均衡，其中高能力雇员选择一种信号，而低能力雇员随机地在高能力雇员选择的信号与另一种信号之间进行选择。特别地，低能力雇员随机地选择与高能力雇员混同或与其分离。

假设高能力雇员选择的信号为e_h，低能力雇员以概率π选择e_h或以概率$(1-\pi)$选择e_L。根据贝叶斯法则，企业观察到信号e_h时的信念为：

$$\begin{aligned}\tilde{P}(H\mid e_h)&=\frac{prob(e_h\mid H)prob(e_h)}{prob(e_h\mid H)prob(H)+prob(e_h\mid L)prob(L)}\\&=\frac{1\cdot P}{1\cdot P+\pi(1-P)}\\&=\frac{P}{P+\pi(1-P)}\end{aligned}\tag{7.27}$$

其中 P 为雇员为高能力类型的先验概率。显然，企业观察到信号e_L时的信念为$\tilde{P}(H\mid e_L)=0$。由式（7.27），因$P+\pi(1-P)\le P+(1-P)=1$，故$\tilde{P}(H\mid e_h)\ge q$，其含义为：由于高能力雇员总选择$e_h$，但低能力雇员只是以概率$\pi$选择$e_h$，故一旦观察到$e_h$被选择，就说明雇员为高能力的概率比先验概率有所提高。当π趋于零时，低能力雇员几乎不会与高能力雇员混同，于是$\tilde{P}(H\mid e_h)$趋于 1，即观察到e_h后几乎可以肯定雇员是高能力的。当π趋于 1 时，低能力雇员几乎总是与高能力雇员混同，故$\tilde{P}(H\mid e_h)$趋于 P。

当低能力雇员选择e_L与高能力雇员分离时，有$\tilde{P}(H\mid e_L)=0$，则有$W(e_L)=y(L,e_L)$，对低能力雇员来说，给定这种工资率，其最优信号为$e^*(L)$。所以，必有$e_L=e^*(L)$。低能力雇员在$e^*(L)$与e_h之间随机选择，据混合博弈最优战略性质即推论 3.2.1 有

$$W^*(L)-C[L,e^*(L)]=W_h-C(L,e_h)\tag{7.28}$$

即他在选择$e^*(L)$与e_h之间无差异，其中$W_h=W(e_h)$。

据式（7.10）有

$$W_h=\frac{P}{P+(1-P)\pi}y(H,e_h)+\frac{(1-P)\pi}{P+(1-P)\pi}y(L,e_h)\tag{7.29}$$

给定e_h，式（7.28）决定一个w_h，若w_h满足

$$y(L,e_h)\le W_h\le y(H,e_h)$$

则式（7.29）决定一个唯一的π，否则不存在杂合均衡。

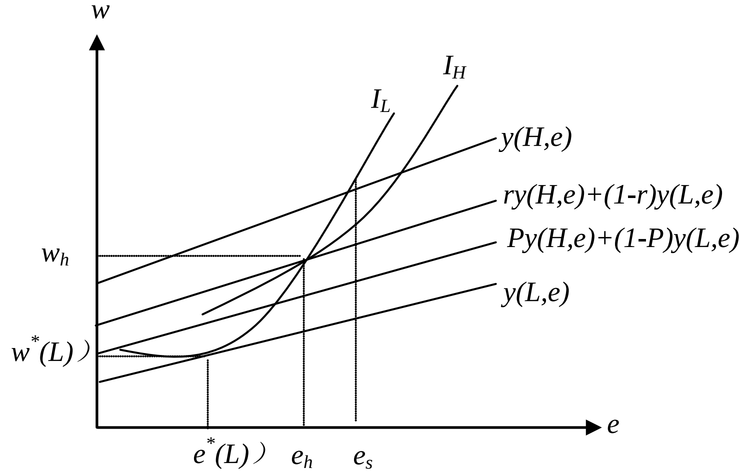


图 7.9　杂合均衡

在图 7.9 中，给定e_h，w_h为式（7.28）的解，(e_h, w_h)处于低能力雇员通过的$[e^*(L), w^*(L)]$的无差异曲线上。

r是$w_h = ry(H, e_h) + (1-r)y(L, e_h)$的解，且$r = \tilde{P}(H \mid e_h)$。据式（7.27），$\pi = \dfrac{P(1-r)}{r(1-P)}$。条件$w_h < y(H, e_h)$等价于$e_h < e_s$，而$e_s$是分离均衡中高能力雇员选择的信号。当$e_h$趋于$e_s$时，$r$趋于 1，故$\pi$趋于零。于是，图 7.8 描述的分离均衡为这里考虑的杂合均衡的极限。

所以，这里考虑的杂合均衡可如下描述：

令

$$\tilde{P}(H \mid e) = \begin{cases} 0 & e < e_h \\ r & e \ge e_h \end{cases}$$

企业战略为

$$W(e) = \begin{cases} y(L, e) & e < e_h \\ ry(H, e) + (1-r)y(L, e) & e \ge e_h \end{cases}$$

对于低能力雇员，在$e < e_h$下的最优信号为$e^*(L)$，且在$e \ge e_h$下的最优信号为e_h。对于高能力雇员，e_h优于任何其他信号。

斯宾塞于 1973 年的论文存在的一个不足是即使教育对于生产率完全没有影响（即即使能力为θ的雇员产出等于$y(\theta)$，与e无关），工资率也可随教育程度的提高而提高（Spence, 1973）。他在 1974 年的论文中，使其论证更加一般化一些，允许产出不仅随能力而提高，还随教育而提高，类似的结论变为工资率随教育而提高的幅度大于教育对生产率的促进可以解释的水平（Spence, 1974）。我们在这里介绍的是这种更加一般的方

法。

那么，从经验考察看，在高校读书时间更长的雇员，其工资率也会更高（平均而言）吗？Mincer（1974）的研究证实了这一预测。这一事实使得我们用在高校读书的时间来表示变量e。在分离均衡中，可以想象能力较低的雇员只读到高中毕业，能力较高的雇员则完成了大学教育。但不幸的是，把e解释为高校教育的时间将引起我们在前面的模型中未涉及到的动态问题，比如在雇员大学一年级毕业时（即在能力较低的雇员根据假设离开高校之后，而在能力较高的雇员根据假设离开学校之前）企业向其开出工资率的可能性。在更加复杂的模型中，雇员可在每一年选择是接受当年所开出的最高工资率出去工作，还是继续念书，第二年再进行选择。Noldeke 和 Van Damme（1990）分析了属于上面类型的更为复杂的博弈模型，并证明如下结论：

（1）存在多个精炼贝叶斯均衡；

（2）通过和我们将在后面运用的再精炼方法非常相近的精炼之后只有一个均衡存留下来；

（3）在（2）中存留的唯一均衡与我们前面分析的博弈通过后面给出的再精炼之后的唯一均衡完全相同。因而在上述分析的博弈中，我们不太严格地把e看成高校教育的时间，因为在更为丰富的博弈中的结果是一致的。

不过，在这里我们又将e的差异解释为学生在学校里的表现上的差异，而不是学生接受高校教育时间的差异，从而绕过上面的动态问题。据此，上述博弈就可以运用于一组高中毕业生（即接受整整 12 年教育的雇员），或一组大学毕业生或 MBA。在这种解释下，e衡量所学课程的门数及种类，在固定的教育期内取得的高分质量和赢得的荣誉等。学费支出（若存在）从而和e不再相关，于是成本函数$c(\theta,e)$量度非货币的（或精神上的）成本；在给定的学校中，能力较低的学生发现要取得好成绩更加困难，同时在一所更具竞争力的学校中要得到相同的分数也更困难。用教育以此种方法作为一个信号也可说明下面的事实：企业雇佣给定学校中最好的毕业生和最好学校的毕业生，并支付他们更高的工资。

7.3.2 公司融资的方式选择

对于一家公司来说，如果它面临一个项目的开展，就需要融资。融资有多种可选择的方式或渠道，如发行股票或发行债券。当然，公司也可以选择寻找内部资金的方式来解决项目的资金投入问题。这被称为公司的资本结构问题。决定公司资本结构选择或融资方式选择的因素有很多，但来自博弈论的逻辑是一个颇有影响的方面。

假定有一家公司面临一个待开展的项目，需要投入资金。公司的盈利能力（即好的企业或差的企业）是企业家的私人信息，并且新项目的收益与企业在开展新项目之前的收益是无法区分的，只有企业的总利润水平是可观察的（当然，可以假设企业家知道新

项目单独的盈利能力，但这种假设除了增加数学上的复杂性外不会带来本质上的不同内容。)。这是一个信号博弈，其中公司和投资者是二个局中人，公司是信号发送者，投资者是信号接收者。公司发送的信号是有多种选择的，我们先来看信号是公司向投资者承诺的股权份额的情形。

为了简化分析，假定公司只有两种类型，即高盈利能力 H 和低盈利能力 L，且 H>L>0（用数量刻画盈利能力）。这里，H 和 L 分别是公司在高盈利或低盈利能力时的利润水平（可以是不确定情形下的期望利润水平）。以下用π表示公司利润水平。设项目需要投入的资金数量为 I，预期收益为 R，而潜在投资者若将数额为 I 的资金投入其它项目中的最大收益为$I(1+r)$，其中r为其它投资渠道的最大回报率（机会成本）。

博弈顺序为：

1．“自然”选择公司的盈利能力，设$\pi=L$的先验概率为 P;

2．公司知道自然的选择π，它向潜在投资者发送信号S，即承诺给投资者的股权份额为S，$0\le S\le 1$；

3．投资者观察到S，然后选择是接受还是拒绝这一合约；

4．如果投资者拒绝合约，则投资者的支付为$I(1+r)$，公司的支付为π；如果投资者接受，则投资者支付为$S(\pi+R)$，公司支付为$(1-S)(\pi+R)$。

这一模型由 Myers 和 Majluf 给出（Myers, S 和 N. Majluf，1984），但他们当时考虑的是一种大公司（有股东和经理层）而不是私人企业（只有唯一一个所有者充当公司总经理）。他们在这篇论文中讨论了股东利益如何影响经理效用的不同假定情形。Dybvig 与 Zender 后来推导出了股东和经理之间的最优合约安排（Dybvig，P.与 J. Zender，1991）。

当投资者接收到合约 S 后，如果他在此信息集上关于公司类型为 L 的后验概率为q，则当且仅当如下条件成立时他会接受S（假设投资者是风险中性的）：

$$S[qL+(1-q)H+R]\ge I(1+r) \tag{7.30}$$

而对公司来说，如果他只有以股权融资与放弃项目两种选择，则它要选择让渡公司的 S 份额股权作为代价为项目融资，当且仅当如下条件成立：

$$S\le \frac{R}{\pi+R} \tag{7.31}$$

下面看看混同均衡条件。在混同均衡中，投资者观察到 S 后的后验概率等于先验概率，即$q=P$。此时，立项约束式（7.31）在$\pi=H$时比$\pi=L$时更加难以满足，式（7.30）和式（7.31）给出的联合约束意味着混同均衡必须满足如下必要条件：

$$\frac{I(1+r)}{PL+(1-P)H+R}\le \frac{R}{H+R} \tag{7.32}$$

当P趋近于零时，上式自然成立，因为已假设有$R>I(1+r)$。但当P趋于 1 时，则上式成立的条件是

$$\frac{I(1+r)}{L+R} \le \frac{R}{H+R}$$

即

$$R - I(1+r) \ge \frac{I(1+r)H}{R} - L \qquad (7.33)$$

即要有足够高的项目收益才会有混同均衡。在式（7.30）中令 $q = P$，得

$S \ge I(1+r)/[PL + (1-P)H + R]$，而若投资者确信有 $\pi = H$（即 $q = 0$），则他将接受更小的权益份额 $S \ge I(1+r)/(H+R)$。这就是说，当 P 足够大时，混同均衡要求更大的 S 才能使投资者接受，因为他担心有较大可能将资金投入一个差的企业，故要求较高的 S 补偿其风险损失。所以，当 P 较大时，高盈利能力企业实际上补贴了低盈利能力企业。我们将在后面的分析中看到，混同均衡所要求高盈利能力企业以更大的权益分额对高盈利能力企业来说可能是非常昂贵的，甚至高到使高盈利能力企业选择放弃这一有盈利前景的新项目。这一结果表明，或者只有当 P 足够小时，才存在混同均衡，这时减少了高盈利能力企业的补贴成本，或者，当式（7.33）成立，此时新项目有足够的利润足以补偿补贴成本，以下将证明这一结论。

当式（7.32）不成立时，则不存在混同均衡，但可证明博弈总存在分离均衡。在分离均衡中，低盈利能力企业发出信号为 $S_L = I(1+r)/(L+R)$，这是投资者愿接受的最低合约水平。高盈利能力企业发出信号 $S_H \prec I(1+r)/(H+R)$。此时投资者会拒绝。这是一个分离均衡。当 $S_L = I(1+r)/(L+R)$，投资者接收到 S 后认为发送者是低盈利能力企业，$q = 1$。当他拒绝时，支付为 $I(1+r)$，而接受时支付为 $S_L(L+R) = I(1+r)$，他在接受与拒绝之间是无差异的。假设此时他选择接受。当投资者接收到 $S_H < I(1+r)/(H+R)$ 时，他知道发送者是高盈利能力企业，$q = 0$。若接受，支付为 $S_H(H+R)$，若拒绝，支付为 $I(1+r) > S_H(H+R)$，故最优选择是拒绝。在非均衡路径上，若 $\frac{I(1+r)}{(H+R)} \le S < I(1+r)/(L+R) = S_L$，信念规定为 $q = 1$，投资者拒绝。

若 $S > I(1+r)/(L+R)$，信念规定为 $q = 1$，投资者接受。给定投资者的信念与选择，低盈利能力企业若选择 $S < I(1+r)/(L+R)$，则投资者拒绝，其支付为 L，小于发送 $S_L = I(1+r)/(L+R)$ 时的支付 $(1-S_L)(L+R)$，因为

$$(1-S_L)(L+R) > \left(1 - \frac{I(1+r)}{(L+R)}\right)(L+R) = L + R - I(1+r) > L$$

若选择 $S > I(1+r)/(L+R) = S_L$，则投资者接受，其支付为 $(1-S)(L+R)$，小于发送 S_L 时的支付 $(1-S_L)(L+R)$。故低盈利能力企业的最优信号为 S_L。

对于高盈利能力企业，当发送 $S \ge \frac{I(1+r)}{(H+R)}$ 时，若 $S < S_L$，则投资者拒绝，支付为 H，若发送 $S < \frac{I(1+r)}{(H+R)}$，则投资者拒绝，支付为 H；若 $S \ge S_L$，则投资者接受，支付为

$(1-S)(H+R)$，但当他发送$S<\dfrac{I(1+r)}{(H+R)}$时，投资者拒绝，支付为 H，此时式（7.32）不成立，有

$$\frac{I(1+r)}{L+R}\geq\frac{I(1+r)}{PL+(1-P)H+R}>\frac{R}{H+R}$$

故

$(1-S)(H+R)\leq(1-S_L)(H+R)=\left(1-\dfrac{I(1+r)}{L+R}\right)(H+R)\leq\left(1-\dfrac{R}{H+R}\right)(H+R)=H$，故发送$S\geq\dfrac{I(1+r)}{(H+R)}$时的支付或者等于 H 或者小于 H，因而最优信号是$S<\dfrac{I(1+r)}{(H+R)}$。这就证明了这一分离均衡的存在。

在这类分离均衡中，投资水平无效率地降低了，因为高盈利企业无法通过股权融资，而低盈利企业却可以通过股权融资。新项目肯定是可以带来利润的，但获得高利润类型的企业却放弃了通过股权融资的投资。这种情形是发送者的信号集无效率的的情形，因为高盈利企业无法将自己与低盈利企业区分开来——对高盈利能力类型有吸引力的融资条件对低盈利能力类型甚至更有吸引力。

这样，当分离均衡出现时，高盈利能力企业就不能选择股权融资方式投资了，这时需要选择其它公式融资。这正是为什么企业在股权融资外还要选择债务融资或内部资金渠道进行投资的原因，即企业的资本结构呈现多元化的特征。

下面，我们考察公司在选择股权融资的同时还可以选择债务融资的情形。

假设企业发出的信号是一份债务合同D，如果企业没有破产，则投资者支付为D，企业支付为$(\pi+R-D)$；如果企业破产，投资者支付为$(\pi+R)$（债权人清理企业资产），企业支付为零。当$L>0$时，总存在混同均衡——两种类型企业均发出相同的信号$D=I(1+r)$，且投资者都接受。这是因为，给定信号$D=I(1+r)$，投资者后验概率等于先验概率，$q=P$。

因为此时企业总不会破产$(\pi+R-D=\pi+R-I(1+r)\geq\pi\geq L>0)$，则接受时投资者支付为$D=I(1+r)$，不接受时支付也为$I(1+r)$，投资者在接受与拒绝之间是无差异的。在已有假定即投资者在此无差异状态下总选择接受，则接受是投资者的最优选择。在非均衡路径上，信号$D\neq I(1+r)$，规定后验概率如下：

$$\tilde{P}(L\mid D)=\begin{cases}1 & D>I(1+r)\\0 & D<I(1+r)\end{cases}$$

故当$D>I(1+r)$时接受，因为接受的支付为$D>I(1+r)$，不接受的支付为$I(1+r)<D$；当$D<I(1+r)$，当然不接受。

给定投资者的上述反应，高盈利能力企业若发出$D\geq I(1+r)$，投资者接受，支付为$(H+R-D)$，若它发出$D<I(1+r)$，投资者不接受，支付为 H，最优选择当然为发出

$D = I(1+r)$；低盈利能力企业若发出 $D \geq I(1+r)$，投资者接受，支付为 $(L+R-D)$，若发出 $D < I(1+r)$，投资者不接受，支付为 L，最优选择为发出 $D = I(1+r)$。

但是，如果 $L < 0$，且使得 $L + R - I(1+r) < 0$，则低盈利能力企业会破产，投资者不会接受这一合约。此时，这一有盈利前景的新项目仍然不能通过债务合同融资[5]。但若 $L > 0$，上述分析表明债务合同可以为新项目融得资金。这就是为何企业有时选择股权融资而有时又选择债权融资，从而形成企业多元化资本结构的原因之一。当 $L + R - I(1+r) < 0$ 时，不可能通过债权融资，如果股权融资也不行，则企业只有选择用内部资金投资了。于是，企业在 $\dfrac{I(1+r)}{PL + (1-P)H + R} > \dfrac{R}{H+R}$ 和 $L + R - I(1+r) < 0$ 时会选择用内部资金投资新项目（因 $R > I(1+r)$，故投资新项目总是盈利的）。如果 R 很大，L 很小（且 $L < 0$），这两个条件显然是可能成立的。

7.3.3 货币政策的动态一致性

在第 5 章中，我们运用子博弈精炼均衡的概念指出货币政策中的动态不一致性，即给定公众预期政府会实行零通胀政策，政府就有动机实行通胀政策。但是，尽管零通胀政策是一个帕累托最优的货币政策，它却不是一个子博弈精炼的均衡。在第 5 章中，我们假定动态博弈是完全信息的。这里，一旦我们引入不完全信息假定，零通胀货币政策就有可能成为精炼贝叶斯均衡，而政府的零通胀政策就可能变成可以置信的承诺了。这是因为，在信息不完全情形，公众可能并不清楚政府在通胀与就业目标之间的替代程度，因而政府可能伪装成不喜欢通胀的类型。当这种伪装行为达成一种混同均衡时，公众就可能相信政府，从而产生较低的通胀率预期。

在第 5 章中，我们假定政府的单阶段支付函数为

$$M(\pi, y) = -c\pi^2 - (y - k\bar{y})^2$$

其中，$c > 0$，$k > 1$ 为常数，π 为政府制造的通胀率，y 为国民收入（总产出），$\bar{y}$ 为充分就业的国民收入。

总产出与通胀率之间的关系由含有公众预期通胀率的 Phillips 曲线刻画：

$$y = \bar{y} + \beta(\pi - \pi^e), \beta > 0$$

其中 β 为常数，π^e 为公众预期的通胀率。

于是有 $M(\pi, y) = M(\pi, .\pi^e) = -c\pi^2 - [(1-k)\bar{y} + \beta(\pi - \pi^e)]^2$　（7.34）

与第 5 章不同的是，这里假定这一动态博弈是二阶段的，第一阶段由公众选择预期通胀率 π_1^e，由政府选择通胀率（货币政策）π_1，第二阶段公众观察到 π_1，选择 π_2^e，政府再选择 π_2。在这种动态博弈中，每一阶段也是一个动态博弈。与第 5 章不同，我们在这里还要给出公众的支付函数，其单阶段支付函数假定为 $-(\pi - \pi^e)^2$，即公众不喜欢预期通胀率与实际通胀率偏差太远[6]。

记 π_t，π_t^e，$t = 1,2$ 分别为第一、二阶段的通胀率和预期通胀率，并假设政府和公众

的贴现因子皆为 1，则政府和公众的支付函数分别为

$$u_g = M(\pi,\pi_1^e) + M(\pi_2,\pi_2^e)$$
$$u_P = -M(\pi_1 - \pi_1^e)^2 - (\pi_2 - \pi_2^e)^2 \quad （7.35）$$

其中 u_g、u_P 分别为政府和公众的支付函数。

在政府的支付函数中的参数 C 反映了政府在零通胀率和充分就业两个宏观经济政策目标之间的替代情况。当 C 较大时，给定同样的通胀率，政府支付较小，即 C 刻画了政府厌恶通胀的程度。这样，当政府为追求充分就业目标而通过制造通胀而使总产出提升时，较大的 C 就意味着政府因此而付出较大的代价。在第 5 章中，C 是局中人的共同知识，而这里，我们假定：C 不再是共同知识，而是政府的私人信息或政府的类型。

设 C 只有两种取值

$$C = \begin{cases} S & \text{强政府(政府对治理通胀的强硬态度)} \\ W & \text{弱政府} \end{cases}$$

$S > W$（因 C 刻画了政府讨厌通胀的程度）

博弈顺序为：

（1）“自然”选择政府的类型 C，$C = W$ 的先验概率为 P；

（2）公众形成通胀率预期 π_1^e；

（3）政府观察到 π_1^e，选择通胀率 π_1；

（4）公众观测到 π_1（不能观察到 C），然后形成通胀率预期 π_2^e；

（5）政府考察到 π_2^e，然后选择通胀率 π_2。

我们从这一两阶段的货币政策博弈中抽象出单阶段信号博弈。发送者的信号为政府对第一阶段通胀率的选择 π_1，接收者的行动为公众对第二阶段通胀率的预期 π_2^e。公众第一阶段的通胀率预期以及政府第二阶段对真实通胀率的选择分别为信号博弈之前及之后的行动。

在信号博弈的单阶段问题中，给定公众的通胀率预期 π^e，政府对 π 的最优选择由如下最大化问题决定：

$$\max_{\pi} M(\pi_1,\pi^e)$$

由式（7.34），该问题的解为

$$\pi_1 = \frac{\beta[(k-1)\bar{y} + \beta\pi^e]}{C + \beta^2} \quad （7.36）$$

在第二阶段，当政府类型为 C，公众预期为 π_2^e，则政府对 π_2 的最优选择为

$$\pi_2^*(\pi_2^e, C) = \frac{\beta[(k-1)\bar{y} + \beta\pi_2^e]}{C + \beta^2} \quad （7.37）$$

如果公众的后验概率为 $\tilde{P}(w \mid \pi_1) = q$，并据此开始第二阶段的博弈，则他们将选择 $\pi_2(q)$，以使下式最大化：

$$-q[\pi_2^*(\pi_2^e,w)-\pi_2^e]^2-(1-q)[\pi_2^*(\pi_2^e,S)-\pi_2^e]^2 \qquad (7.38)$$

下面考虑混同均衡，此时两种类型选择的第一阶段通胀率相同，不妨用π^*表示。假设公众具有理性预期，于是，公众第一阶段的预期为$\pi_1^e=\pi^*$。

在均衡路径上，公众观察到π^*，后验概率为$\widetilde{P}(w\,|\,\pi^e)=P$，然后开始第二阶段博弈，并形成预期$\pi_2^e(P)$，类型为 C 的政府对给定预期选择最优的第二阶段通胀率。在理性预期假定下有$\pi_2^*[\pi_2^e(P),C]$，然后博弈结束。在非均衡路径上，据式（7.38）计算相应的非均衡路径上的行为。当政府在第一阶段选择$\pi_1=\pi^*$，第二阶段选择$\pi_2^*(\pi_2^e,c)$，且π_2^e是最大化如下函数：

$$-P[\pi_2^*(\pi_2^e,w)-\pi_2^e]^2-(1-P)[\pi_2^*(\pi_2^e,S)-\pi_2^e]^2 \qquad (7.39)$$

其中$\pi_2^*(\pi_2^e,w)$和$\pi_2^*(\pi_2^e,S)$由式（7.37）决定。

政府的支付为$M(\pi^*,\pi^*)+M(\pi_2^*(\pi_2,c),\pi_2^e)$。

当政府在第一阶段发出信号$\pi_1\neq\pi^*$时，规定第二阶段公众的后验概率为$\widetilde{P}(w\,|\,\pi_1)=1$，则公众在第二阶段的最优预期通胀率为最大化，$-[\pi_2^*(\pi_2^e,w)-\pi_2^e]^2$的$\pi_2^e$，它满足

$$\pi_2^*(\pi_2^e,w)=\pi_2^e$$

由式（7.37）得：$\pi_2^e=\dfrac{\beta(k-1)\bar{y}}{c}$ （7.40）

此时政府的支付为$M(\pi_1,\pi^*)+M\left(\dfrac{\beta(k-1)\bar{y}}{c},\dfrac{\beta(k-1)\bar{y}}{c}\right)$

当下述条件成立时，π^*是混同均衡信号。

$$M(\pi^*,\pi^*)+M[(\pi_2^*(\pi_2^e,c),\pi_2^e]\geq M(\pi_1,\pi^*)+M\left(\frac{\beta(k-1)\bar{y}}{c},\frac{\beta(k-1)\bar{y}}{c}\right) \qquad (7.41)$$

其中π_2^e是最大化式（7.39）的解，$\pi_2^*(\pi_2^e,c)$由式（7.37）决定。

$c=s,w$， $\pi_1\in(-\infty,+\infty)$

当我们将第一阶段视为一个足够长的时期，而第二阶段是政府任期的末期，则通过混同均衡，即使是弱政府，也可能通过与强政府的混同而带来强政府的声誉从而获得一个较低的通胀率。只要式（7.41）能给出较小的π^*。此时，在第一阶段，就达成一种低通胀率的均衡[7]。

下面看看分离均衡，此时在第一阶段，强政府和弱政府分别发出信号π_s和π_w，于是$\pi_1^e=P\pi_w+(1-P)\pi_s$。第二阶段在均衡路径上，若公众观察到信号$\pi_1=\pi_w$，则$\pi_2^e$由最大化如下函数决定：

$$-[\pi_2^*(\pi_2^e,w)-\pi_2^e]^2$$

记其解为$\pi_2^e(1)$。当公众观察到π_s时，则π_2^e由最大化如下函数决定：

$$-[\pi_2^*(\pi_2^e,s)-\pi_2^e]^2$$

记其解为$\pi_2^e(0)$。第二阶段，强政府和弱政府分别选择$\pi_2^e(\pi_2^e(0),s)$和$\pi_2^*(\pi_2^e(1),w)$。

我们可以规定非均衡路径上的公众后验概率为

$\tilde{P}(w\mid\pi_1)=1$，$\pi_1\neq\pi_s,\pi_w$

此时公众在第二阶段的预期通胀率为 $\pi_2^e(1)$。

当下述条件成立时，π_s,π_w 为分离均衡信号[8]：

$$M(\pi_s,\pi_1^e)+M[\pi_2^*(\pi_2^e(0),S),\pi_2^e(0)]$$
$$\geq M(\pi_1,\pi_1^e)+M[\pi_2^*(\pi_2^e(1),S),\pi_2^e(1)]$$
$$\pi_1\neq\pi_s$$
$$M(\pi_w,\pi_1^e)+M[\pi_2^*(\pi_2^e(1),w),\pi_2^e(1)]$$
$$\geq M(\pi_1,\pi_1^e)+M[\pi_2^*(\pi_2^e(1),w),\pi_2^e(1)]$$
$$\pi_1\neq\pi_w$$

这里，排除了弱政府在第一阶段伪装强政府从而选择 π_s，进而为诱使公众第二阶段的预期为 $\pi_2^e(0)$，并在其后选择 $[\pi_2^*(\pi_2^e(0),w]$ 使博弈结束的可能。这种可能是由于即使 π_s 较低，以至于弱政府有些不情愿，但它会使 $\pi_2^e(0)$ 非常低，使之可从第二阶段的预料外通货膨胀 $[\pi_2^*(\pi_2^e(0),w]$—$\pi_2^e(0)$ 之中获得巨大收益。在分离均衡中，强政府选择的第一阶段通胀率必须足够低，使弱政府没有动机去伪装成强政府，即使在第二阶段可获得预料外通胀的好处。一般情况下，这一约束使得 π_s 低于强政府在完全信息下将会选择的通胀率水平，这就好象前述 7.3.1 中 Spence 模型的分离均衡中，高能力雇员将过度投资于教育一样，是为了显示自己的类型而产生的信号成本。

与 7.3.1 中的 Spence 模型类似，这一博弈存在许多混同、杂合及分离均衡[9]。

7.3.4 Milgrom–Roberts 垄断限价模型

Milgrom 和 Roberts (1982)的垄断限制性定价模型（limit pricing model）是信号博弈在产业组织理论中的第一个应用，它也是信号博弈中的一个经典例子。该模型试图解释这样一个现象：

在现实中，有些垄断企业实际上并未按微观经济学给出的最优垄断价（即由边际成本等于边际收益决定的价格）定价，而是低于这种垄断价格。人们对该现象的一种十分自然的可能解释是：垄断企业为了阻止其它企业进入，故意将价格定得偏低，利润薄一些，使其它潜在企业看到无利可图，不进入该行业，从而避免竞争。

但这种解释存在的缺陷是：如果垄断企业的成本函数是进入者已知的，那么，这种低价是不可置信的承诺，因为若垄断性的在位者是高成本，对进入者来说无论在位者如何制造假象，它也会进入（因它知道在位者是高成本）。Milgrom 和 Roberts 给出的解释自然就是：可能是因为进入者不知道在位者的成本情况，而在位者就通过用低价来告诉进入者自己是低成本(即使在位者可能实际上是高成本)，进入是无利可图的。

模型

局中人：两阶段博弈　企业 1—在位者；企业 2—进入者。

企业 1 的战略空间：第一阶段选价格 p_1；第二阶段若进入者不进入，则选短期垄断价，若进入，则选古诺博弈均衡价。

企业 2 的战略空间：第二阶段决定是否进入，若进入，与在位者进行古诺博弈。

不完全信息假定：企业 1 有两个可能类型：$\begin{cases}\text{高成本}H,\text{先验概率为}u(H)\\ \text{低成本}L,\text{先验概率为}1-u(H)\end{cases}$

记企业 1 选价格 p_1 时的短期垄断利润为 $M_1^{\theta}(p_1),\theta=H,L$。用 p_m^{θ} 表示类型 θ 的最优垄断价格，$M_1^{\theta}=M_1^{\theta}(p_m^{\theta})$ 表示最大短期垄断利润，其中 $p_m^H>p_m^L,M_1^H<M_1^L$。假定 $M_1^{\theta}(p_1)$ 是严格凹函数。[10]

用 D_1^{θ} 和 D_2^{θ} 分别表示当企业 1 为类型 θ 时，企业 1 和 2 在第二阶段的利润（D_2^{θ} 是扣除进入成本后的净利润）。

设 $D_2^H>0>D_2^L$（这样分析才有意义）。δ 是共同的贴现因子。显然，在完全信息下，当只当 1 是高成本时，2 才会进入。因 1 希望保持垄断地位（因 $M_1^{\theta}>D_1^{\theta}$），他想让 2 知自己是低成本，但即使他真的是低成本，他也无法直接达到此目的。一个间接办法是定一个低价 p_1^L，但有时即使 1 是高成本也会如此定一个低价将 2 阻止在市场外，只要一定条件满足即可。

分离均衡

在分离均衡中，高、低成本在位者的身份被进入者准确知道，故高成本在位者必选 $p_1^H=p_m^H$（短期最优价），且进入者必进入，故高成本在位者的总利润为 $(M_1^H+\delta D_1^H)$。

设此时低成本在位者选 p_1^L。对此 p_1^L，高成本在位者之所以不去模仿（此时是分离均衡），必有条件

$$M_1^H(p_1^L)+\delta M_1^H\le M_1^H+\delta D_1^H$$

即模仿低成本在位者在分离均衡中可阻止进入者进入（分离均衡中，当在位者是低成本时，进入者不敢进入），但其总利润还不如直接暴露自己高成本身份引致进入的利润来得大。记该条件为：

$$M_1^H+\delta D_1^H\ge M_1^H(p_1^L)+\delta M_1^H \tag{7.42}$$

或

$$M_1^H-M_1^H(p_1^L)\ge\delta(M_1^H-D_1^H) \tag{7.43}$$

低成本在位者选 p_1^L 并能阻止进入（分离均衡中），其总利润为 $M_1^L(p_1^L)+\delta M_1^L$，若他选短期垄断价，其总利润为 $M_1^L+\delta x$，在最坏情况下，进入者进入，$x=D_1^L$，一般有 $x\ge D_1^L$，故必有 $M_1^L+\delta x\ge M_1^L+\delta D_1^L$。

他所以选 p_1^L 而不是短期垄断价，必然是有

$$M_1^L(p_1^L)+\delta M_1^L\ge M_1^L+\delta D_1^L \tag{7.44}$$

或

$$M_1^L - M_1^L(p_1^L) \le \delta(M_1^L - D_1^L) \tag{7.45}$$

条件（7.42）和（7.43）的含义

p_1^L 是低成本在位者的这样一种价格，它低得使高成本模仿无意义，因模仿尽管可阻止进入，却在总利润上比不模仿直接暴露自己高成本身份（从而引致进入）还低；它又足够高，使得低成本在位者自己的总利润比直接选短期垄断价而且造成进入者进入时带来的总利润高。

注意：这里原则上并未排除 $p_1^L = p_m^L$ 的情形，即低成本在位者直接选短期垄断价时，高成本在位者不敢模仿。但这里假定不存在 $p_1^L = p_m^L$ 的分离均衡，即 $p_1^L \ne p_m^L$。这一假定即是说，p_m^L 不是条件（7.42）和（7.44）中 p_1^L 的共同解，但 p_m^L 显然是条件（7.44）的解（满足条件（7.44）），故 p_m^L 必然不会是条件（7.42）（(7.43)）的解，于是有

$$M_1^H - M_1^H(p_m^L) < \delta(M_1^H - D_1^H) \tag{7.46}$$

我们给出一种 Spence–Mirrlees**（分离条件）条件**（SM **条件**）

$$\frac{\partial}{\partial p_1}\left[M_1^H(p_1) - M_1^L(p_1)\right] > 0$$

或

$$\frac{\partial M_1^H(p_1)}{\partial p_1} > \frac{\partial M_1^L(p_1)}{\partial p_1}$$

当（SM）条件成立时，可证明存在区间 $\left[\tilde{\tilde{p}}, \tilde{p}\right]$，使任何 $p_1^L \in \left[\tilde{\tilde{p}}, \tilde{p}\right]$ 都满足条件（7.42）和（7.44），即构成一个分离均衡价格。另可证条件（7.46）意味着 $\tilde{p} < p_m^L$，故为了得到分离均衡，低成本在位者需要选一个足够低的价格（低于短期垄断价 p_m^L）使得高成本在位者模仿的成本太高。

SM 条件是说：改变价格对不同类型企业的利润的影响是不同的。高成本企业比低成本企业更愿意选高价。该条件一般是满足的。假如边际成本为常数，分别为：

C^H 和 C^L，$C^H > C^L$，需求函数为 $Q(P_1)$，则

$$\frac{\partial M_1^H(P_1)}{\partial P_1} = \frac{\partial}{\partial P_1}\left[\left(P_1 - C^H\right)Q\left(P_1\right)\right] = Q(P_1) + \left(P_1 - C^H\right)\frac{\partial Q(P_1)}{\partial P_1}$$

$$\frac{\partial M_1^L(P_1)}{\partial P_1} = \frac{\partial}{\partial P_1}\left[\left(P_1 - C^L\right)Q\left(P_1\right)\right] = Q(P_1) + \left(P_1 - C^L\right)\frac{\partial Q(P_1)}{\partial P_1}$$

因 $C^H > C^L$，$Q'(P_1) < 0$，故 SM 条件满足。

若成本类型是连续分布的，则 SM 条件变为

$$\frac{\partial^2 M(P,C)}{\partial P \partial C} = \frac{\partial}{\partial P \partial C}\left[\left(P - C\right)Q(P)\right] = -\frac{\partial Q(P)}{\partial P} > 0$$

SM 条件使曲线 $Y = M_1^L - M_1^L(P_1^L)$ 与 $Y = M_1^H - M_1^H(P_1^L)$ 在 (P_1^L, Y) 空间只交叉一次，见图 7.10。

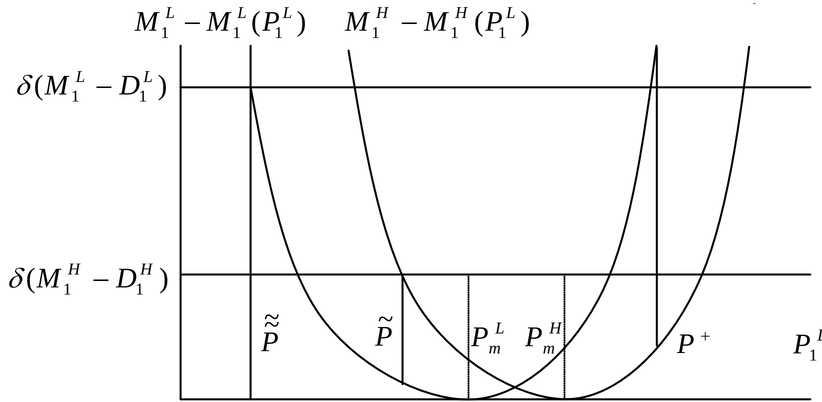


图 7.10　垄断限价模型的分离均衡与混同均衡定价区域

显然，所有 $P_1^L \in \left[\tilde{\tilde{P}}, \tilde{P}\right]$ 都满足分离均衡必要条件（7.43）和（7.45）。下面证明 $P_1^L \in \left[\tilde{\tilde{P}}, \tilde{P}\right]$ 还是分离均衡的充分条件。

任给 $P_1^L \in \left[\tilde{\tilde{P}}, \tilde{P}\right]$，将证在位者给定下述战略是一个最优战略（对应于某一后验概率）。

在位者战略 $\begin{cases} \text{高成本选} P_1 = P_m^H \\ \text{低成本选} P_1 = P_1^L \end{cases}$

后验概率：均衡路径，$\tilde{u}\left(H \middle| P_1 = P_m^H\right) = 1$，$\tilde{u}\left(H \middle| P_1 = P_1^L\right) = 0$（即 $\tilde{u}\left(L \middle| P_1 = P_1^L\right) = 1$），不难验证，它们满足贝叶斯法则。

非均衡路径：为了保证在位者战略是最优的，非均衡路径上的后验概率需要限制在某些范围，一个简单的办法是：令 $\tilde{u}\left(H \middle| P_1 \neq P_m^H, P_1^L\right) = 1$，即观察到非均衡路径上的行动后都认为是高成本，从而选进入，下面证明在位者战略在此后验概率上是最优的。

若是高成本在位者，他选 P_m^H 得总利润 $M_1^H + \delta D_1^H$，他选 P_1^L 得总利润 $M_1^H(P_1^L) + \delta M_1^H$，由条件（7.42）知前者大于后者，故不会选 P_1^L。他若选其它 $P_1 \neq P_m^H, P_1^L$，则由后验概率知进入者会进入，其总利润不会大于 $M_1^H + \delta D_1^H$，

故他选 P_m^H 为最优。

若是低成本在位者，他选 P_1^L 得总利润 $M_1^L(P_1^L) + \delta M_1^L$，他若选 P_m^H 得总利润 $M_1^L(P_m^H) + \delta D_1^L \leq M_1^L + \delta D_1^L$，由条件（7.44）知前者大于后者，故他不会选 P_m^H。

他若选其它 $P_1 \neq P_m^H, P_1^L$，则给定后验概率意味着进入者进入，他的总利润为 $M_1^L(P_1) + \delta D_1^L \leq M_1^L + \delta D_1^L$，条件（7.44）也意味着总利润没有他选 P_1^L 时大。

故 P_1^L 是低成本在位者的最优选择。

原在位者战略在给定后验概率下是最优的，故是一个精炼贝叶斯均衡。

这里只是给出一个特定的后验概率，与此在位者战略构成精炼贝叶斯均衡的后验概率还有许多。

我们得到连续的无穷多个分离均衡，即高成本在位者选 P_m^H，低成本在位者选任何的 $P_1^L \in \left[\tilde{\tilde{P}}, \tilde{P}\right]$，$\tilde{u}\left(H\middle|P_1^L\right)=0$，$\tilde{u}\left(H\middle|P_1 \neq P_1^L\right)=1$。到底哪一个是最合理的均衡呢？恐怕应是 $P_1^L = \tilde{P}$，因它使低成本在位者以最低的成本将自己与高成本在位者区分开来，且不影响高成本在位者的利润（高成本在位者总是选 P_m^H），故 $P_1^L = \tilde{P}$ 是“帕累托最优”分离均衡价格。

只要 $u(H) > 0$，这样的连续分离均衡就存在，但当 $u(H)$=0，低成本在位者就会选短期垄断价 P_m^L（因 $u(H)$=0 使贝叶斯法则对先验概率无影响），这意味着，信息结构的小小变化就会导致均衡结果的很大不同：只要进入者认为在位者是高成本的先验概率 $u(H)$>0，低成本在位者就不得不非连续地（即从 P_m^L 一下下跳到 $\tilde{P} < P_m^L$）降低价格以显示自己是低成本。就是说，不完全信息博弈对信息结构是非常敏感的。

混同均衡

在混同均衡中，进入者从先行动者行动中得不到新的信息，高、低成本在位者又选择相同的价格，则进入者必不敢进入，因为若进入者进入，则高、低成本在位者也知进入者会进入的条件下，他们会分别在进入者给定要进入情况下，选价格 P_m^H 和 P_m^L，这就违反了混同均衡假定（高、低成本在位者选同一价格）。故进入者得不到新信息下，认为在位者是高、低成本的后验概率分别等于先验概率 u 和(1–u)，他不敢进入说明有

$$uD_2^H + (1-u)D_2^L < 0 \quad (7.47)$$

设 P_1 是一个混同均衡价格。

对于低成本在位者，他选 P_1 的总利润为 $M_1^L(P_1) + \delta M_1^L$（在混同均衡中），他选 P_m^L 的总利润为 $M_1^L + \delta x$，而 $x \geq D_1^L$，他在混同均衡中必选 P_1，故有

$$M_1^L(P_1) + \delta M_1^L \geq M_1^L + \delta D_1^L \quad (7.48)$$

或

$$M_1^L - M_1^L(P_1) \leq \delta(M_1^L - D_1^L) \quad (7.49)$$

这就是前述条件（7.48）（只不过这里 P_1 是混同均衡价，那里的 P_1^L 是分离均衡价）。

对于高成本在位者，他选混同均衡价 P_1 的利润为 $M_1^H(P_1) + \delta M_1^H$（混同均衡中），他选 P_m^H 的总利润为 $M_1^H + \delta x, x \geq D_1^H$，故有

$$M_1^H(P_1) + \delta M_1^H \geq M_1^H + \delta D_1^H$$

或

$$M_1^H - M_1^H(P_1) \leq \delta(M_1^H - D_1^H) \quad (7.50)$$

可证明：在条件（7.46）满足下，在 P_m^L 周围存在一个区间 $\left[\tilde{P}, P^+\right]$（见前图 7.10），使所有 $P_1 \in \left[\tilde{P}_1, P^+\right]$ 都是式（7.49）和（7.50）的共同解。故 $\left[\tilde{P}, P^+\right]$ 是混同均衡的必要条件。下面证明 $\left[\tilde{P}, P^+\right]$ 还是充分条件。

在位者战略：高、低成本在位者都选 P_1， $P_1 \in \left[\tilde{P}, P^+\right]$，后验概率：均衡路径，$\tilde{u}\left(H \middle| P_1\right) \equiv u(H)$，不进入；非均衡路径，$\tilde{u}\left(H \middle| P \neq P_1\right) = 1$（$\tilde{u}\left(L \middle| P \neq P_1\right) = 0$），进入。

对于高成本在位者，选 P_1 时总利润为 $M_1^H(P_1) + \delta M_1^H$，选其它价 $P_1^{'}$ 时总利润为 $M_1^H(P_1^{'}) + \delta D_1^H \leq M_1^H + \delta D_1^H$，由条件（7.50）知前者大于后者，故选 P_1 最优。

对于低成本在位者，选 P_1 时总利润为 $M_1^L(P_1) + \delta M_1^L$，选其它价 $P_1^{'}$ 时总利润为 $M_1^L(P_1^{'}) + \delta D_1^L \leq M_1^L + \delta D_1^L$，由条件（7.48）知前者大于后者，故选 P_1 最优。

注意，还有其它一些后验概率与给定 $P_1 \in \left[\tilde{P}, P^+\right]$ 构成精炼贝叶斯均衡。

我们得到连续的无穷多个混同均衡。现代博弈论研究中正在寻找额外的理由来去掉其中大多数均衡，从而使合理的均衡保持少数（最好只有一个），如在分离均衡中，帕累托最优价为 $P_1^L = \tilde{P}$，在混同均衡中，$P_1 = P_m^L$ 是帕累托优于（P_1, P_m^L）的，P_m^H 帕累托优于（P_m^H, P^+），既然“混同”的真正意义是高成本在位者把自己混同于低成本在位者从而阻止进入，他们都选大于 P_m^L 的价是很不合理的。

7.3.5 用负债比例显示企业质量：Ross 模型

自 1950 年代以来，经济学家一直在探讨一个重要问题，即企业的资本结构对于企业价值有何影响。

企业价值指企业的市场价值。

米勒和莫迪蒂安尼曾证明了一个著名的“MM 定理”，指出企业价值与企业资本结构无关。

企业资本结构指企业资本中债务与股权的比例。

但是，现实中企业总是不会随意地设计其资本结构，而是力图使其资本结构达到某一特定比例。

这说明企业价值与其资本结构是有关的。因为企业经理的福利与企业价值有关。

直到目前为止，许多人一直在研究“反 MM 定理”，即企业价值与资本结构有关的具体机制，该领域现在也是金融经济学中的热点之一。

但是，这类研究的许多定理都只是通过改变“MM 定理”中原有的假定出发来获得企业价值与资本结构有关的结论，而下面将给出的一个结果却与此相异，是从信号博弈出发，认为资本结构之所以与企业价值有关，是由于资本结构作为一个信号，企业向公众传达了有关企业质量的信息，从而影响公众（市场）对企业价值的评价。

这一结果（即资本结构的信号博弈理论）是该领域内最有影响的理论之一。

该理论由 Ross（1977）提出，其中心思想是：

企业内部经理与市场上的外部投资者之间存在信息不对称，资本结构通过传递内部信息对企业的市场价值发生影响。

下面我们介绍 Ross 模型如下：

假定：企业经理知企业在第二阶段中的市场价值的真实分布函数，而投资者不知道（在第二阶段，企业市场价值由企业的由公众预期的利润分布决定，企业经理比外部投资者拥有更多的关于企业赢利能力的信息。在第一阶段，企业经理力图通过发出信号使公众对其企业价值给予正确的评价，达到分离均衡，但第二阶段的企业市场价值将会受到第一阶段行为的影响，如债务对后一阶段有影响，企业经理必须兼顾第一阶段的显示企业质量与其作为代价对后一阶段经营的影响，后面将用加权平均两个阶段企业市场价值（后一阶段是期望值）作为企业经理的效用函数来表达这一兼顾性。

假定：在第二阶段，企业市场价值（由其获取利润的能力决定）的分布函数是根据一阶随机占优排序的，即企业愈好（质量愈高），其市场价值分布区域的上限愈高（获高利润的概率愈高或市场价值愈可能在高水平处出现）。

此博弈行动顺序为：

第一阶段，企业经理通过发出信号，让金融市场投资人知其是高质量企业，从而踊跃购买其股票和债券，因而给予其股票和债券以较高价格，赋予企业较高价值。

第二阶段，企业用从金融市场上融来的资金进行经营，此时，企业的市场价值由其预期利润的资本化决定（与土地价值决定类似），而由于利润率是随机的，故为一随机变量，但高质量企业获高利润的概率较大。

对于企业经理来说，企业市场价值（股票和债券价值）愈高，其个人福利就愈大，相反，当企业破产时，经理受到惩罚（失去工作，名誉损失等）。

经理通过企业的负债比例（负债占总资产的比例）向投资者传递企业获利能力的信息，而投资者将较高负债率看作是企业高质量的表现，原因是：

$$\text{破产概率}\begin{cases}\text{与企业质量负相关}\\ \text{与企业负债率正相关}\end{cases}$$

倘若企业是低质量的，其破产概率（在给定负债率下）较大，如果还提高负债率，则破产概率更大。

反之，高质量企业的破产概率小，提高负债率（给定企业质量下）尽管会提高破产概率，但破产概率不会变得很大的，故这是低质量企业不敢模仿的。

高质量企业可以通过提高负债率的办法来显示自己的高质量。

表现为：越是好的企业，负债率就越高。

以下分析模型。

假定：两个时期，两个局中人，即企业经理和投资者。

令 v 为企业在第二时期的市场价值(预期利润流的资本化)，v 在区间[0, θ]上均匀分布，经理知θ，投资者只知θ的概率分布（先验概率）$\mu(\theta)$（θ是企业的类型），θ愈大，企业质量愈高。在第一阶段，经理先选择负债水平 D，投资者观察到 D，从而通过金融市场上的行动决定企业市场价值v_0。在第二阶段，企业用金融市场融来的资金实现利润（受到第一阶段负债 D 的不利影响，D 愈大，第二阶段破产概率愈大）。

下面给出经理的效用函数。

假定：经理效用函数是第一、二阶段企业市场价值的加权平均（由前讨论，他要兼顾 2 个阶段的企业表现）。

在第一阶段，企业市场价值为$V_0(D)$；在第二阶段，对经理来说，记企业的“价值”v'是一个随机变量（获利润受多种因素影响），一方面，即使企业不破产，其价值也受到影响利润的多种因素的影响，另一方面，当破产时，企业“价值”v 为 $v–L$（L 为破产惩罚），即将企业以 v 卖了以还债，但受到惩罚 L，实得为 $v–L$（可能为负数）。

因当 $v<D$ 时，就出现破产，故有

第二阶段“价值”$v' = \begin{cases} v & v \ge D \\ v-L & v < D \end{cases}$

经理将第二阶段“价值” v'的期望水平作为指标，从而进入效用函数进行加权。

“价值” v'的期望水平$=\int_0^\theta$"价值"f(价值)d(价值)$=\int_0^\theta "v'" f(v')dv'$

（其中，f(价值)$=f(v')$为“价值” v'的分布密度。）

$=\int_0^D (v-L)f(\text{价值})dv+\int_D^\theta vf(\text{价值})dv$

$=\int_0^D (v-L)f(v)dv+\int_D^\theta vf(v)dv$

（因 $f(v)$ 为$[0,\theta]$上均匀分布，故$\int_0^\theta f(v)dv = f(\text{常数})\theta = 1$，所以有 $f=\frac{1}{\theta}$）

$=\frac{\int_0^D vdv}{\theta}-\frac{L\int_0^D dv}{\theta}+\frac{\int_D^\theta vdv}{\theta}$

$=\frac{\int_0^\theta vdv}{\theta}-\frac{LD}{\theta}$

$=\frac{\frac{1}{2}\theta^2}{\theta}-\frac{LD}{\theta}$

$=\frac{1}{2}\theta-\frac{LD}{\theta}$

效用（支付）函数为

$$u(D,V_0(D),\theta)=(1-\gamma)V_0(D)+\gamma\left(\frac{1}{2}\theta-\frac{LD}{\theta}\right)$$

其中γ为加权权数。

因第二阶段企业市场价值期望水平为$\int_0^\theta vfdv=\frac{\int_0^\theta vdv}{\theta}=\frac{\theta}{2}$，故$\frac{\theta}{2}$为第二阶段企业的期望价值。

$\frac{D}{\theta}$是企业破产概率，即

企业破产概率$=P(v<D)=\int_0^D fdv=\frac{D}{\theta}$

假定$D\le\theta$，若有$D>\theta$，则令破产概率$\frac{D}{\theta}=1$（概率不能大于 1，这里潜在假定企业没有其他资产可以用来抵债）。

设经理选 D 时，投资者观察到 D 后，认为企业属于类型θ的概率（后验概率）为$\tilde{u}(\theta/D)$，则他认为企业类型的期望值为$\tilde{\theta}(D)=\int_0^\infty\theta\tilde{u}(\theta\mid D)d\theta,\tilde{u}(\theta\mid D)$由贝叶斯公式给出：

$$\tilde{u}(\theta\mid D)=\frac{P(D\mid\theta)u(\theta)}{\int_0^\infty P(D\mid\theta)u(\theta)d\theta}$$

企业市场价值（期望水平）$=\int_0^{\tilde{\theta}(D)}vfdv=\frac{\frac{1}{2}\tilde{\theta}^2(D)}{\tilde{\theta}(D)}=\frac{\tilde{\theta}(D)}{2}$

下面只考虑分离均衡。

因有$\frac{\partial^2u(D.V_0(D),\theta)}{\partial D\partial\theta}=\frac{\gamma L}{\theta^2}>0$

Spence–Mirrlees 条件成立（质量(θ)愈高的企业，愈不害怕负债）。

存在分离均衡的单一信号区间。

下面计算经理的最佳信号：

$$\frac{\partial u}{\partial D}=\frac{1}{2}(1-\gamma)\frac{\partial\tilde{\theta}(D)}{\partial D}-\gamma L\frac{1}{\theta}=0$$

因在分离均衡下，投资人从 D 正确地推断出θ，即当$D(\theta)$是类型θ的企业经理的最优选择，则有$\tilde{\theta}(D(\theta))=\theta$

$\frac{d\tilde{\theta}(D(\theta))}{d\theta}=1$（事实上，分离均衡下有$P(\theta\mid D(\theta))=1$，$P(\theta'\mid D(\theta))=0$，$\theta'\neq\theta$，故

$$\tilde{\theta}(D)=\int_0^\infty\theta'\tilde{u}(\theta\mid D)d\theta'=\int_0^\infty\theta'P(\theta'\mid D(\theta))d\theta'=\int_0^\infty\theta'\delta(\theta'-\theta)d\theta=\theta$$

δ为克朗内克函数）。

$$\frac{d\tilde{\theta}}{dD}\cdot\frac{dD(\theta)}{d\theta}=1$$

$$\frac{d\tilde{\theta}}{dD}=\left(\frac{dD}{d\theta}\right)^{-1}$$

代入上述一阶条件：

$$\frac{1}{2}(1-\gamma)\left(\frac{dD}{d\theta}\right)^{-1}-\frac{\gamma L}{\theta}=0$$

$$2\gamma L\frac{dD}{d\theta}-(1-\gamma)\theta=0$$

解此微分方程：

$D(\theta)=\left(\frac{1-\gamma}{4\gamma L}\right)\theta^2+C$，$C$ 为积分常数。

于是得到经理的均衡战略，即类型 θ 的企业经理选择负债水平 $D(\theta)$。

$$\theta=\left[(D(\theta)-C)\frac{4\gamma L}{(1-\gamma)}\right]^{\frac{1}{2}}$$

因“理性预期”性的观察（分离均衡下能准确识别身份），$\tilde{\theta}=\theta$。

$$\tilde{\theta}=\left[(D-C)\frac{4\gamma L}{(1-\gamma)}\right]^{\frac{1}{2}}$$

代入 $V_0=\frac{\tilde{\theta}}{2}$

$$V_0=\left[(D-C)\frac{\gamma L}{(1-\gamma)}\right]^{\frac{1}{2}}$$

在此博弈中（给定 C，对应一个博弈），D 的取值范围为 $[C,\infty)$，而对于 $[C,+\infty)$ 中任一个 D，都有某类型 θ 的企业经理选 D。

任何 D 都是均衡路径，故无非均衡路径。

对每一观察到的 D，由贝叶斯法则给出一个后验概率 $\tilde{u}\left(\theta^{-1}(D)\mid D\right)=1$ 和 $\tilde{u}\left(\overline{\theta}\neq\theta^{-1}(D)\mid D\right)=0$。

此精炼均衡意味着，愈是高质量的企业，负债水平（及负债率）就愈高，尽管投资者不能直接观察到企业的质量，但他们通过观察企业的负债率来判断企业的质量，从而正确地给企业定价。

非对称信息逼使愈是好的企业承担愈高的负债率，因在完全信息下，企业的负债水平与企业价值无关（由 MM 定理）。因为增加负债提高了企业破产的概率，经理为信息的非对称性付出了成本，这种成本在完全信息下是不存在的。

为了节省这笔成本，企业经理可以向投资者披露内部信息。

这就是现实中看到的上市公司经理有兴趣向金融市场投资者披露内部信息的原因，因为这样使经理用更廉价的方法让投资者知道了企业的真实价值。

从此均衡战略还可看出，若经理愈不重视企业的市场价值，企业的负债率就愈低。因负债的唯一目的是向外部投资者传递信息，若经理不重视企业市场价值，他就没有必要向市场传递信息，从而没有必要举债和承担破产惩罚。

7.4 完全非完善信息两阶段博弈

7.4.1 完全非完美信息两阶段博弈的定义

我们在本章中通过 Harsanyi 转换将不完全信息动态博弈转化为完全但不完美信息动态博弈，这里所说的完全但非完美信息两阶段博弈就是其中的一种重要类型。

有许多模型都可用以下相同的结构加以处理，即两阶段博弈，下一阶段开始前局中人可观察到前面所有阶段的行动，且每一阶段中存在着同时行动（这种同时行动意味着博弈包含了不完美信息）。

定义 7.4　满足以下条件特征的博弈称为完全非完美信息两阶段博弈：

1．局中人 1 和 2 同时从各自的行动空间 A_1 和 A_2 中选择行动 a_1 和 a_2。

2. 局中人 3 和 4 观察到第一阶段的结果 (a_1, a_2) 然后同时从各自的行动空间 A_3 和 A_4 中选择行动 a_3 和 a_4。

3．局中人获得支付 $u_i(a_1, a_2, a_3, a_4), i = 1,2,3,4$

许多经济学问题都符合以上条件，还有很多经济问题可通过把以上条件稍加改动而建立模型，比如增加局中人人数或者允许同一局中人（在一个以上的阶段）多次选择行动。也可以允许少于四个的局中人：在一些应用中，局中人 3 和 4 就是 1 和 2；还有的则不存在局中人 2 或 4。

7.4.2 完全非完美信息两阶段博弈的子博弈精炼纳什均衡

求解完全非完美信息两阶段博弈的精炼均衡可以沿用逆向归纳法的思路，但这里从博弈的最后阶段逆向推导的第一步就包含了求解一个真正的博弈（给定第一阶段结果时，局中人 3 和 4 在第二阶段同时行动的博弈）。

假设　为使问题简化，以下假设对第一阶段博弈每一个可能结果 (a_1, a_2)，其后（局中人 3 和 4 之间的）第二阶段博弈有唯一的纳什均衡 $(a_3^*(a_1, a_2), a_4^*(a_1, a_2))$。

如果局中人 1 和 2 预测到局中人 3 和 4 在第二阶段的行动将由 $(a_3^*(a_1, a_2), a_4^*(a_1, a_2))$ 给出，则局中人 1 和 2 在第一阶段的问题就可用以下的同时行动博弈表示：

1．局中人 1 和 2 同时从各自的行动空间 A_1 和 A_2 中选择行动 a_1 和 a_2；

2．局中人的支付为$u_i(a_1,a_2,a_3^*(a_1,a_2),a_4^*(a_1,a_2)),i=1,2$

假定(a_1^*,a_2^*)为以上同时行动博弈唯一的纳什均衡，我们称$(a_1^*,a_2^*,a_3^*(a_1^*,a_2^*),a_4^*(a_1^*,a_2^*))$为这一两阶段博弈的子博弈精炼解。

此解与完全且完美博弈中的逆向归纳解在性质是一致的，并且与后者有着类似的优点和不足。如果局中人 3 和 4 威胁在后面的第二阶段博弈中，他们将不选择纳什均衡下的行动，局中人 1 和 2 是不会相信的，因为当博弈确实进行到第二阶段时局中人 3 和 4 中至少有一个人不愿把威胁变为现实（恰好是因为它不是第二阶段博弈的纳什均衡）。

另一方面，假设局中人 1 就是局中人 3，并且局中人 1 在第一阶段并不选择a_1^*，局中人 4 就会重新考虑局中人 3（即局中人 1）在第二阶段将会选择$a_3^*(a_1,a_2)$的假定。

例 7.2：银行挤兑模型

局中人 1、2 在银行存入一笔存款 D，银行已将这些存款投入一个长期项目。如果在该项目到期前银行被迫对投资者变现，则共可收回$2r$（假定$D>r>D/2$)。不过，如果银行允许投资项目到期，则项目共可取得$2R$（假定$R>D$)。

投资者可在两个日期中从银行提款：

日期 1 在银行的投资项目到期之前，日期 2 则在到期之后。假设贴现因子为 1。

如果两个投资者都在日期 1 提款，则每人可得r，博弈结束。

如果只有一个投资者在日期 1 提款，他可得到 D（因未到期，银行至多归还给其原有投资 D)，另一人得到$2r-D$，博弈结束。

如果两人都不在日期 1 提款，则项目结束后投资者在日期 2 进行提款决策。如果两个投资者在日期 2 提款，则每人得到 R，博弈结束。如果只有一个投资者在日期 2 提款，则他得到$2R-D$（投资者尽量多地索取，留给其他人可接受的 D 数量的回报），另一人得到 D，博弈结束。

最后，如果在日期 2 两个投资者都不提款，则银行向每个投资者返还 R，博弈结束。

表 7.1 给出了分别在日期 1、2 提款的博弈支付矩阵。

表 7.1 银行挤兑的支付矩阵

1 \ 2	提款	不提
提款	r,r	$D,2r-D$
不提	$2r-D,D$	下一阶段

a. 日期 1

1 \ 2	提款	不提
提款	R,R	$2R-D,D$
不提	$D,2R-D$	R,R

b. 日期 2

图 7.11 给出了博弈树。

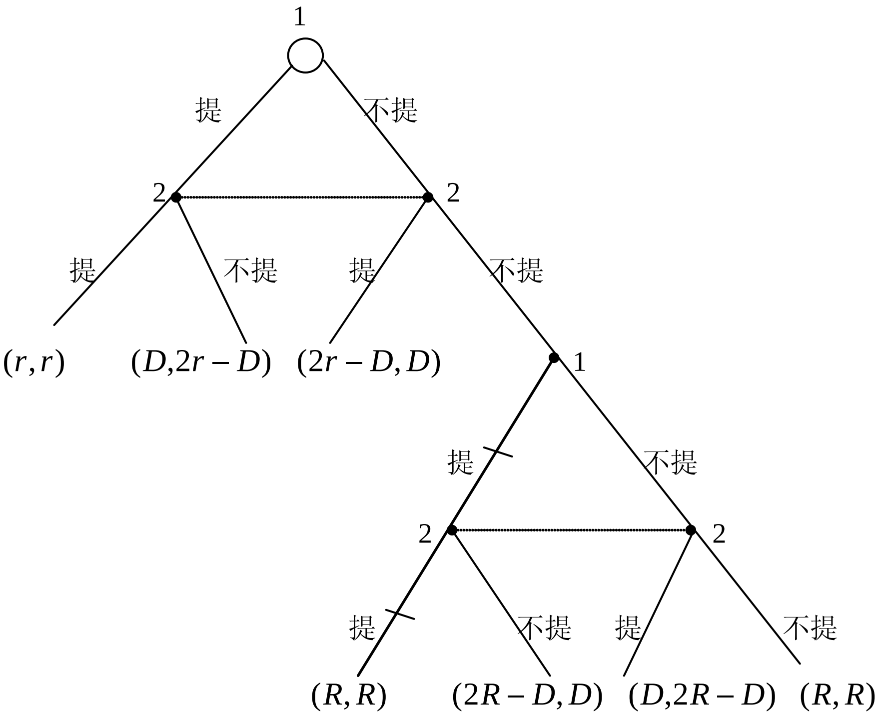


图 7.11　银行挤兑博弈的博弈树

沿用逆向归纳法的思路，有如下推理：

在第二阶段，因 $R>D$，故 $2R-D>R$

对于 2 来说，无论在倒数第一个信息集上的哪一个决策结上，“提”总是严格优于“不提”。

因 $R>D$，故此时 1 的最优选择是“提”。

第二阶段唯一的纳什均衡为：（提款、提款），支付为 (R,R)

因贴现因子为 1，故可直接用这一支付代入日期 1 的博弈矩阵在双方都不提款时的情况，见表 7.2。

表 7.2　将第二阶段博弈结果代入第一阶段时的支付矩阵

		2：提款	2：不提
1	提款	$\underline{r},\underline{r}$	$D,2r-D$
	不提	$2r-D,D$	$\underline{R},\underline{R}$

因$r < D$，故$2r - D < r$，所以在第一阶段中有两个纯战略纳什均衡：

（1）（提款，提款），支付为(r, r)

（2）（不提，不提），支付为(R, R)

第一个均衡结果就是对银行的一次挤提。

由奇数定理知还有一个混合战略博弈均衡存在。

令γ为 2“提”的概率，θ为 1“提”的概率，v_1(提)和v_1(不提)分别是局中人选择“提”和“不提”的期望支付。则有

v_1(提)= $r\gamma + D(1-\gamma)$

v_1(不提)= $(2r - D)\gamma + R(1-\gamma)$

因均衡时有v_1(提)= v_1(不提)（据推论 3.2.1）

故$r\gamma + D(1-\gamma) = (2r - D)\gamma + R(1-\gamma)$

得$\gamma = \dfrac{R - D}{R - \gamma}$

由对称性知$\theta = \gamma = \dfrac{R - D}{R - \gamma}$

均衡$\left\{\left(\dfrac{R-D}{R-r}, \dfrac{D-R}{R-r}\right), \left(\dfrac{R-D}{R-r}, \dfrac{D-r}{R-r}\right)\right\}$

在三个均衡中，何者实际会出现呢？

倘若存在一个信号，使局中人之一认为某均衡会出现，他就会选相应的均衡战略。

其他人发现他选该均衡中的战略，也就会选相应的均衡战略。

于是，该均衡就真的会出现。当然，即使局中人并非根据什么信号而自行选择某个均衡战略，也会因此而导致这个均衡的真正出现，这被称为“自我实现”(Self fulfilling)。

例 7.3 关税和国际市场的不完全竞争

WTO 即世界贸易组织是一个众所周知的国际性自由贸易组织，其宗旨是促进不同国家之间的自由贸易。自由贸易会给所有国家带来帕累托改进，这是古典经济学家就已讲清楚的道理。但为什么还需要一个组织来促进不同国家之间的自由贸易呢？该例给出的模型说明了所有国家之间实行自由贸易是一种帕累托最优但它并不是一种子博弈精炼纳什均衡，因而不会自动出现。WTO 起的作用至少从理论上来说是通过设计特定的规则使最优的状况作为精炼均衡。

局中人 1、2 是两个完全相同的国家，分别用$i = 1,2$表示。

每个国家有一个政府负责确定关税税率，有一个企业制造产品供给本国的消费者及出口，以及一群消费者在国内市场购买本国企业或外国企业的产品。

设i的市场上总供给量为Q_i，市场需求曲线为：

$P_i(Q_i) = a - Q_i$

局中人i的企业（后称企业i）为国内市场生产产量h_i，出口e_i，则

$Q_i = h_i + e_j$

设企业的边际成本为常数c，且假定固定成本为零。故企业i生产的总成本为：

$c_j(h_i, e_i) = c(h_i + e_i)$

产品出口时企业要支付关税：若政府j制定的关税税率为t_j，则企业i向j国家出口e_i必须支付关税$t_j e_i$给政府j。于是，第i个局中人的纯战略为关税税率t_i。

博弈行动顺序

(1)两国政府同时选择关税税率t_1和t_2；

(2)企业观察到关税税率，并同时选择其提供国内消费和出口的产量(h_1, e_1)和(h_2, e_2)；

(3)企业i的支付为其利润，政府i的支付为本国的总福利，其中国家i的支付是总福利，总福利是国家i的消费者享受的消费者剩余，企业i赚取的利润及政府i从企业j收取的关税收入之和。

局中人i的企业利润为：

$\pi_i(t_i, t_j, h_i, e_i, h_j, e_j) = [a-(h_i+e_j)]h_i + [a-(e_i+h_j)]e_i - c(h_j+e_i) - t_j e_i$

在市场需求曲线为$P_i(Q_i) = a - Q_i$下，当市场i的销售量为Q_i时，消费者剩余为$\frac{1}{2}Q_i^2$（见图 7.12）。

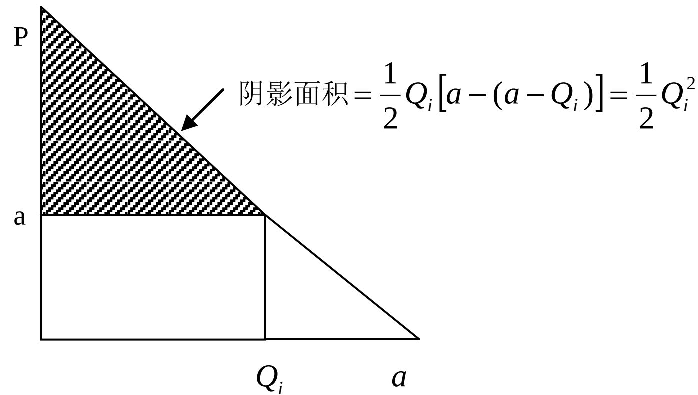


图 7.12　消费者剩余

故总福利函数为：

$$w_i(t_i, t_j, h_i, e_i, h_j, e_j) = \frac{1}{2}Q_i^2 + \pi(t_i, t_j, h_i, e_i, h_j, e_j) + t_i e_j$$

假设政府i选定的税率分别为t_1和t_2，若$(h_1^*, e_1^*, h_2^*, e_2^*)$为企业 1 和企业 2 的（两市场）博弈的纳什均衡，则对每一个企业i，(h_i^*, e_i^*)必须满足

$$\max_{h_i \ge e_i} \pi(t_i, t_j, h_i, e_i, h_j^*, e_j^*)$$

因为$\pi_i(t_i, t_j, h_i, e_i, h_j^*, e_j^*)$可以表示为在市场$i$的利润与在市场$j$的利润之和，而企业$i$在市场$i$的利润只是$h_i$和$e_j^*$的函数，在市场$j$的利润又只是$e_i, h_i^*$和$t_j$的函数，企业$i$在两市场的最优化问题就可简单地折分为一对问题，在每个市场上分别求解：

h_i^*必满足

$\max\limits_{h_i \ge 0} h_i[a-(h_i+e_j^*)-c]$，即在国内市场出售产量的利润极大化

且e_i^*必满足

$\max\limits_{e_i \ge 0} e_i[a-(e_i+h_j^*)-c]-t_j e_i$，即出口利润极大化

一阶条件

$$\frac{d\left\{h_i\left[a-(h_i+e_j^*)-c\right]\right\}}{dh_i}=0$$

$$a-(h_i+e_j^*)-c-h_i=0$$

$$h_i^*=\frac{a-c-e_j^*}{2} \text{（假设} e_j^* \le a-c \text{）} \qquad (7.51)$$

$$\frac{d\left\{e_i\left[a-(e_i+h_j^*)-c\right]-t_j e_i\right\}}{de_i}=0$$

$$a-(h_i+e_j^*)-c-e_i-t_j=0$$

$$e_i^*=\frac{a-c-t_j-h_j^*}{2} \text{（假设} h_j^* \le a-c-t_j \text{）} \qquad (7.52)$$

（以下的结果表明，式（7.51）、（7.52）中的假设都是满足的）

式（7.51）和（7.52）是企业$i=1,2$的反应函数，共有四个方程，求解得：

$$h_i^*=\frac{a-c+t_i}{3} \text{ 且 } e_i^*=\frac{a-c-2t_j}{3} \qquad (7.53)$$

如果没有政府（关税为零），则两企业进行的是古诺博弈，均衡产出都为$\frac{a-c}{3}$，但这一结果是基于对称的边际成本推出的。

而式（7.48）的均衡结果与之不同的是，政府对关税的选择使企业的边际成本不再对称。譬如在市场i，企业的边际成本为c，但企业j的边际成本则为$c+t_i$。由于企业j的成本较高，它意愿的产出也相对较低。但若企业j要降低产出，市场出清价格又会相应提高，于是企业i又倾向于提高产出。

在这种情况下，企业j的产量就又会降低。

结果就是：在均衡下，h_i^*随t_i的提高而上升，e_j^*随t_i的提高而（以更快的速度）下降（见式（7.48））。

在解出了政府选定关税时，其后第二阶段两企业博弈的结果之后，我们可把第一阶段政府间的互动决策表示为以下的同时行动博弈：

首先，政府同时选择关税率t_1和t_2。

第二，政府i的支付为$w_i(t_i,t_j,h_1^*,e_1^*,h_2^*,e_2^*)$，$i=1,2$，这里$h_i^*$和$e_i^*$是式（7.48）所表示的$t_i$和$t_j$的函数。下面来求解这一政府间博弈的纳什均衡。

因为$h_i^*=h_i^*(t_i)$，$e_i^*=e_i^*(t_j)$，故令

$w_i^*(t_i,t_j)=w_i(t_i,t_j,h_1^*,e_1^*,h_2^*,e_2^*)$

它是政府i选择关税t_i，政府j选择关税t_j，企业i和j按式（7.48）中的纳什均衡选择行动时政府i的支付。

设(t_1^*,t_2^*)是这一政府间博弈的纳什均衡，则对每一个i，t_i^*必满足

$$\max_{t_i\geq 0} w_i^*(t_i,t_j^*) \tag{7.54}$$

由$w_i=\dfrac{1}{2}Q_i^2+\pi_i+t_ie_j$

由式（7.48）可求得$Q_i=h_i+e_j$

解出式（7.49）的解为

$t_i^*=\dfrac{a-c}{3}$，$i=1,2$

代入式（7.44）得

$h_i^*=\dfrac{4(a-c)}{9}$且$e_i^*=\dfrac{a-c}{9}$

关税博弈的子博弈精炼解为

$$\left\{t_1^*=t_2^*=\frac{a-c}{3},h_1^*=h_2^*=\frac{4(a-c)}{9},e_1^*=e_2^*=\frac{a-c}{9}\right\}$$

该均衡是低效率的。此时，每一市场上的总销量

$Q_i=\dfrac{5(a-c)}{9}$，但当$t_i=t_j=0$时，由

$Q_i=h_i+e_j$

由式（7.48)，得

$Q_i=\dfrac{a-c}{3}+\dfrac{a-c}{3}=\dfrac{2(a-c)}{3}$

它正好是古诺博弈结果。

此时，市场i的消费者剩余$=\dfrac{1}{2}\times\dfrac{4(a-c)^2}{9}=\dfrac{2(a-c)^2}{9}$

而在博弈均衡时，市场i的消费者剩余$=\dfrac{1}{2}\times\dfrac{25(a-c)^2}{81}=\dfrac{25(a-c)^2}{2\times 81}$

$<\dfrac{2(a-c)^2}{9}$

事实上，零关税是社会最优选择，因$t_1=t_2=0$是下式的解

$\max\limits_{t_1,t_2\geq 0} w_1^*(t_1,t_2)+w_2^*(t_2,t_1)$（不难证明）

于是，政府就有动因签订一个相互承诺零关税税率的协定，这就是促进自由贸易的WTO形成的原理。

但这种承诺若无对违反协议适当的惩罚，则是不可置信的，任一国都有动机私下违反协议。

因在第二阶段，企业i和j将按式（7.44）给出的纳什均衡行动，故政府在第一阶段的互动决策就是令$t_i = \frac{a-c}{3} > 0$。

例 7.4　工作竞赛（Lazear和Rosen（1981）模型）

设有为同一老板工作的两个工人，工人$i(i=2,3)$生产的产出$y_i = e_i + \varepsilon_i$，其中$e_i$为努力程度，$\varepsilon_i$是随机扰动项。

博弈行动的顺序：

(1)两个工人同时选择非负的努力$e_i \geq 0$；

(2)随机扰动项ε_1和ε_2相互独立，并服从期望值为0，密度函数为$f(\varepsilon)$的概率分布；

(3)工人的产出可以预测，但各自选择的努力水平无法预测，从而工人的工资可以决定于各人的产出，却无法（直接）取决于其努力水平。

假设老板为激励工人努力工作，而在他们中间开展工作竞赛。工作竞赛的优胜者（即产出水平较高的工人）获得的工资率为w_H；失败者的工资为w_L。工人获得工资水平w并付出努力程度e时的支付为$u(w,.e) = w - g(e)$，其中，$g(e)$表示努力工作带来的负效用，是递增的凸函数（$g'(e) > 0$且$g''(e) > 0$）。老板的支付为$y_1 + y_2 - w_H - w_L$。下面用前述二阶段重复博弈来说明该博弈过程。局中人1——老板，其行动a_1是选工作竞赛中的工资水平w_H和w_L，这里不存在局中人2。

两个工人是局中人3和4，他们观测第一阶段选定的工资水平，然后同时选行动a_3和a_4，具体地说就是选定的努力程序e_1和e_2（后面还将考虑另一种可能性，即对老板选定的工资水平，工人们不愿意参与工作竞赛，即转而寻找另外的工作机会）。

最后，局中人各自的支付如前所给出。由于产出（并因此而使工资）不只是局中人行动的函数，而且同时还受随机扰动因素ε_1和ε_2的影响，下面用局中人的期望支付进行分析。

假定老板已选了工资水平w_H和w_L，若一对努力水平(e_1^*, e_2^*)是第二阶段两个工人博弈的纳什均衡，则对每个i，e_i^*必使工人的期望工资减去努力带来的负效用后的净支付最大。

即e_i^*必满足

$$\max_{e_i \geq 0} w_H prob\{y_i(e_i) > y_j(e_j^*)\} + w_L prob\{y_i(e_i) \leq y_j(e_j^*)\} - g(e_i) \quad (7.55)$$

$$= (w_H - w_L) prob\{y_i(e_i) > y_j(e_j^*)\} + w_L - g(e_i)$$

其中$y_i(e_i) = e_i + \varepsilon_i$

一阶条件为

$$(w_H - w_L)\frac{\partial prob\left[y_i(e_i) > y_j(e_j^*)\right]}{\partial e_i} = g'(e_i) \tag{7.56}$$

意即：工人i选努力程度e_i，从而使得额外努力的边际负效用$g'(e_i)$等于增加努力的边际效益，后者又等于对优胜者的奖励工资$(w_H - w_L)$乘以因努力程度提高而使获胜概率的增加。

由贝叶斯法则

$$P(A) = \sum_j P(A/B_j)P(B_j)$$

得

$$prob\{y_i(e_i) > y_j(e_j^*)\} = prob\{\varepsilon_i > e_j^* + \varepsilon_j - e_i\}$$

$$= \int_{\varepsilon_j} prob\{\varepsilon_i > (e_j^* + \varepsilon_j - e_i) \mid \varepsilon_j\} f(\varepsilon_j) d\varepsilon_j$$

$$= \int_{\varepsilon_j} \left[1 - F(e_j^* - e_i + \varepsilon_j)\right] f(\varepsilon_j) d\varepsilon_j$$

其中f是概率密度函数，F是分布函数。

一阶条件（7.56）可化为

$$(w_H - w_L)\int_{\varepsilon_j} f(e_j^* - e_i + \varepsilon_j) f(\varepsilon_j) d\varepsilon_j = g'(e_i)$$

因对称性，故有$e_1^* = e_2^* = e^*$，故

$$(w_H - w_L)\int_{\varepsilon_j} f(\varepsilon_j)^2 d\varepsilon_j = g'(e^*) \tag{7.57}$$

因$g(e)$是凸函数，故$g'(e^*) > 0$

优胜者获得的奖励越高（即$w_H - w_L$愈大），就会激发更大的努力（即$g'(e^*)$愈大时，e^*愈大），这与直觉是一致的。

另一方面，在同样奖励水平下，对产出的随机扰动因素愈大。

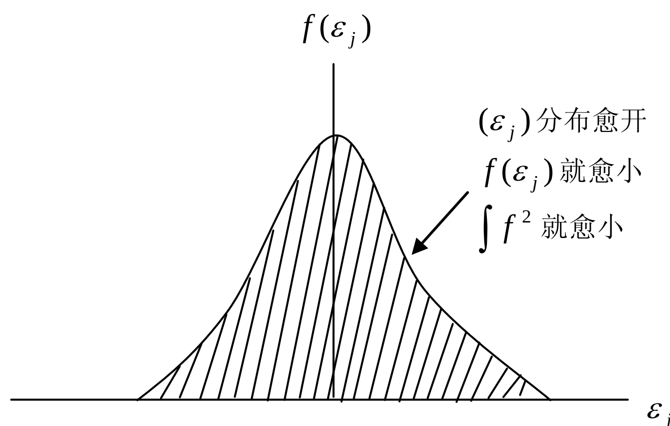


图 7.13　扰动项分布

$\int_{\varepsilon_j} f(\varepsilon_j)^2 d\varepsilon_j$ 就愈小，就愈不值得努力工作，因为这时工作竞赛的最终结果在很大

程度上是决定于运气，而非努力程度。

假如，当ε服从方差为σ^2的正态分布时，则有

$$\int_{\varepsilon_j} f(\varepsilon_j)^2 d\varepsilon_j = \frac{1}{2\sigma\sqrt{\pi}}$$

它随标准差σ的增加而下降，即e^*的确随σ的增加而降低。

再从后往前分析博弈的第一阶段。

假定工人们同意参加工作竞赛（而不是去另谋高就），他们对给定的w_H和w_L的反应，将会是式（7.57）给出的对称纳什均衡战略。

同时假定工人可寻求其他就业机会，得到的效用为U_a。

因在对称纳什均衡中每个工人在竞赛中获得优胜的概率为$\frac{1}{2}$（即$prob\left[y_i(e^*) > y_j(e^*)\right] = \frac{1}{2}$），若老板要使工人有动力参加工作竞赛，则他必须选择满足下式的工资水平

$$\frac{1}{2}w_H + \frac{1}{2}w_L - g(e^*) \ge U_a \quad \text{（这是参与约束）} \qquad (7.58)$$

假设U_a足够低，以致于老板愿意激励工人参加竞赛，则老板会在式（7.51）的参与约束条件下，选择使自己期望支付$2e^* - w_H - w_L$最大化的工资水平。

由于在最优条件下，式（7.58）中的等号成立（老板压低工资使其支付最大）。（式（7.57）中只含$(w_H - w_L)$，与w_L的绝对量无关）。

$$w_L = 2U_a + 2g(e^*) - w_H \qquad (7.59)$$

故期望利润成为$2e^* - 2U_a - 2g(e^*)$

故老板要考虑的问题就是使$e^* - g(e^*)$最大化。

他此时选择的工资水平应使得与之相应的e^*满足这一条件，从而最优选择下的努力程度满足一阶条件：

$$g'(e^*) = 1 \qquad (7.60)$$

代入式（7.60），则意味着最优激励$w_H - w_L$满足

$$(w_H - w_L)\int_{\varepsilon_j} f(\varepsilon_j)^2 d\varepsilon_j = 1 \qquad (7.61)$$

由式（7.59）、（7.60）和（7.61）就可解出w_H和w_L。

7.5 声誉模型

7.5.1 KMRW 声誉模型

我们在第五章中已发现：只要重复博弈次数是有限的，则合作性的子博弈精炼均衡总是不可能出现的。

这种情况似乎与我们在现实中所观察到的许多合作现象的存在相矛盾。因为现实中

的重复博弈一般总是有限次的。事实上，Axelrod（1981）所做的实验表明，即使在有限次重复博弈中，合作行为也是频繁出现的，特别是在距博弈结束仍比较远的阶段。

被称为博弈论“四人帮”的 Kreps、Milgrom、Roberts 和 Wilson（1982）认为：之所以存在前述结论，原因是存在两个假定，即理性人是共同知识假定和信息完全假定。

他们构造了一个著名的“声誉模型”（reputation model），在有限次重复博弈中引入不完全信息和理性人非共同知识的假定，发现存在合作型子博弈精炼均衡解，从而解开了这个悖论。[11]

我们以“囚徒困境”有限次重复博弈为例说明声誉模型的基本思想:

当局中人 2 不知道局中人 1 是否理性（他以概率 P 认为局中人 1 是非理性的，以概率（1-P）认为局中人 1 是理性的），而局中人 1 也知道这一事实时，局中人 1 可以在博弈的前面的若干重复阶段中“伪装”成非理性人，让局中人 2 在前面的每个阶段博弈后不能获取局中人 1 是否理性的进一步信息，局中人 2 当然也知道局中人 1 会如此“伪装”，故他不会在观察到局中人 1 的“伪装性行动选择”后就简单地认为局中人 1 是非理性的，而只是不能从中获取局中人 1 理性程度的进一步信息，从而对局中人 1 的理性程度判断仍为最初的先验信息，即以 P 的概率为非理性，（1-P）的概率为理性。

局中人 1 以这种战略诱使局中人 2 在博弈的大部分过程中总是以 P>0 的概率猜想他是非理性人，从而认为在各个阶段博弈中认为局中人 1 有可能不选“坦白”而选“抵赖”，从而诱使局中人 2 在各个阶段博弈中可能选“抵赖”，这有利于局中人 1 的利益。到了最后一次阶段博弈，任何局中人的最优选择都是选“坦白”（没有了以后的博弈，最后一次阶段博弈实际上是一个单阶段博弈，唯一的均衡是大家都“坦白”）。既然最后一次博弈的结果是确定了的，之前的选择对其无影响，那么，局中人 1 在倒数第二阶段博弈的最佳选择是：当局中人 2 还被蒙住在鼓里以为他可能是非理性的（从而认为他会选“抵赖”而自己也选“抵赖”）时候，自己选“坦白”获得额外好处（这时暴露自己的真实类型不会对其后的博弈结果产生影响，故没有其他损失）。

局中人 1 在博弈的开始直到最后倒数第二阶段博弈的过程中，一直隐藏自己的真实类型，通过牺牲短期利益而“装傻”，为以后的行为创造了一个“声誉”，让局中人 2 总不能获得关于其类型的除先验信息以外的额外信息，从而诱使局中人 2 也选择合作行为（这对双方都是有利的），直到最后（此时声誉已用尽）“露出庐山真面目”。

这种思想还可以扩展到局中人 1 对局中人 2 的理性程度也具有不完全信息的情形，即局中人 1 和 2 的理性人假定都不是共同知识。

模型

以下构造有限次囚徒困境重复博弈的声誉模型。其中有二个模型，第一个模型只假定关于局中人 1 的理性人假定不是共同知识，这二个模型假定关于两个局中人的理性人假定都不是共同知识。

假定 对于局中人 2 来说，他认为局中人 1 有两种潜在的类型，即理性的和非理性的，先验概率分别为（1-P）和 P。但局中人 1 认为局中人 2 只有一个类型，即是理性人。

这里，局中人 2 对局中人 1 的类型只拥有不完全信息。

对于理性人来说，在有限次重复博弈中只会选择不合作行为，即总是选“坦白”。但对于非理性人来说，其可能的选择是任意可假定的（对应于任意可选择的支付函数）。譬如，可假定对非理性人来说，可假定他总是选择“抵赖”，从而局中人 2 可在每次阶段博弈都选“坦白”而获利，但这种非理性的局中人 1 也实在太傻了，即他不论对方如何损害他，他也总是“仇将恩报”。在这种情形，只有单方面的一厢情愿合作行为，合作均衡是不出现的。

我们作出其它的更具合作性一些的关于非理性人行为选择的假定，即能带来合作均衡的假定。

假定 假定非理性的囚徒 1 在以下的博弈中中采用这种战略，即“针锋相对”（tit-for-tat）：开始选“抵赖”，然后在 t 阶段选择囚徒 2 在 $t-1$ 阶段所选择的行动（即“你抵赖我抵赖，你坦白我也坦白”）。

博弈的行动顺序如下：[12]

1．自然选择囚徒 1 的类型；囚徒 1 知道自己的类型，但囚徒 2 只知道囚徒 1 属于理性人的概率为（1-P），是非理性的概率为 P；

2．囚徒 1 和 2 进行第一阶段博弈；

3．他们观测到第一阶段博弈的结果后，进行第二阶段博弈；观测到第二阶段博弈结果后，进行第三阶段博弈等等；

4．理性囚徒 1 和囚徒 2 的支付分别是其各阶段博弈的支付的贴现值之和（为简单化，这里假定贴现因子 $\delta=1$）。

当囚徒 2 和囚徒 1 是理性人时，囚徒 2 必然会在每个阶段选择“坦白”，而理性的囚徒 1 也知道这种结果，故会在各个阶段中选择“坦白”，从而唯一的子博弈精炼均衡是他们在各个阶段都同时选“坦白”，这是我们在前面已经知道的事实。但是，当囚徒 2 认为囚徒 1 为非理性的概率 P 充分大时，尽管他选“抵赖”仍有可能遭遇到囚徒 1 实际上为理性人（概率为 1-P）从而选“坦白”的灾难性结果，但当 P 充分大时，这种可能性较小，从而他在各个阶段都选“抵赖”时遇到囚徒 1 为非理性的（从而选“抵赖”）的可能性较大，所获得的支付增加较大，以至于可以抵消因猜错囚徒 1 的理性属性所可能带来的潜在损失，此时，囚徒 2 会在倒数第一阶段以前的各个阶段都选“抵赖”（最后阶段必然选“坦白”）。

理性的囚徒 1（囚徒 1 实际上是理性的，这是博弈论对局中人是理性人的假定的直

接结果，只不过这里假定理性人假定对局中人来说不是共同知识，即局中人 2 不知道局中人 1 到底是否理性人）也知道囚徒 2 的上述选择（当 P 充分大时），故为了获得合作解，他也会“将计就计”，故意伪装成非理性人，即在倒数第二阶段之前的各阶段都选“抵赖”。

局中人 2 当然也知道理性囚徒 1 的上述选择，但当他预测到对方的合作行为时，不知道囚徒 1 到底是因为其为非理性而选择了“抵赖”呢，还是他是理性人因上述盘算后作出的选择。

因此，理性囚徒 1 通过选“抵赖”而成功地隐藏了有关他理性程度的进一步信息，让囚徒 2 关于囚徒 1 的理性属性判断仍然停留在最初的先验概率（囚徒 1 是非理性的概率为 P）上，从而诱使囚徒 2 在下个阶段仍然按前面各阶段同样的计算作出“抵赖”的选择，而自己也选“抵赖”得到合作解。

直到倒数第二阶段，理性囚徒 1 露出庐山真面目，选“坦白”而出卖囚徒 2，此时囚徒 2 在下个阶段知道了囚徒 1 是理性的，但为时已晚，因为最后一个阶段的选择是确定的（选“坦白”），不受之前选择的影响。

这就是声誉模型的要义：即囚徒 1 与囚徒 2 可能（P 充分大时）达成一个合作（在倒数第二阶段之前的各阶段博弈中）型的子博弈精炼贝叶斯均衡。

我们在这里关于非理性囚徒 1 所以选“针锋相对”战略而不选各阶段都“坦白”的战略的假定，并非是说囚徒 1 可能是“非理性人”，而是指囚徒 2 认为囚徒 1 的支付函数并非是支付矩阵所给出的那种，而可能是一种特殊的其他支付函数。

定理 7.1

在有限 T 次不完全信息的“囚徒困境”重复博弈中，若 $P \geq 2/9$，则以下战略组合构成一个子博弈精炼贝叶斯均衡。

非理性囚徒 1：选“针锋相对”战略；

理性囚徒 1：第一阶段选“抵赖”，一直选“抵赖”直到 T-2 阶段，第 T-1 和 T 阶段选“坦白”。

囚徒 2：开始选“抵赖”，若观测到囚徒 1 选了“坦白”，贝叶斯法则给出的后验概率为 $\tilde{P} = 0$，选“坦白”；若预测到囚徒 1 选“抵赖”，后验概率 P=P，选“抵赖”；最后，在 T 阶段选“坦白”。

证明：对 T 用数学归纳法证明该定理。

显然，要使定理 1 有意义，必须有 $T \geq 3$。

阶段博弈的支付矩阵由表 2.2 给出为：

表 2.2 囚徒困境博弈

		囚徒 2 坦白	囚徒 2 抵赖
囚徒 1	坦白	-8,-8	0,-10
	抵赖	-10,0	-1,-1

记 C 为“坦白”（confess），D 为“抵赖”（Deny）。

我们选证明定理 1 在 T=3 时成立。

首先，给定囚徒 1（理性和非理性）的战略，我们证明囚徒 2 的战略是最优的。

事实上，给定囚徒 1（理性和非理性）的战略。囚徒 2 所可能选择的四种战略（因 T=3 阶段的选择总是 C，是确定了的）分别为：

a. D、D、C

b. D、C、C

c. C、D、C

d. C、C、C

下面将证明战略 a 是最优的。

分别计算囚徒 2 在各个战略下的预期支付水平 u_a、u_b、u_c 和 u_d：

a 战略下

$$\begin{aligned} u_a &= u(D,D,C) \\ &= [P(-1)+(1-P)(-1)]+[P(-1)+(1-P)(-10)]+[P\cdot 0+(1-P)(-8)] \\ &= 17P-19 \end{aligned}$$

	t=1	t=2	t=3
非理性囚徒 1	D	D	D
理性囚徒 1	D	C	C
囚徒 2	D	D	C

b 战略下：

$$\begin{aligned} u_b &= u(D,C,C) \\ &= [P(-1)+(1-P)(-1)]+[P\cdot 0+(1-P)(-8)]+[P\cdot(-8)+(1-P)(-8)] \\ &= 8P-17 \end{aligned}$$

	t=1	t=2	t=3
非理性囚徒 1	D	D	C
理性囚徒 1	D	C	C
囚徒 2	D	C	C

c 战略下：

$u_c = u(C, D, C)$

$= [P \cdot 0 + (1-P) \cdot 0] + [P(-10) + (1-P)(-10)] + [P \cdot 0 + (1-P)(-8)]$

$= 8P - 18$

	t=1	t=2	t=3
非理性囚徒 1	D	C	D
理性囚徒 1	D	C	C
囚徒 2	C	D	C

d 战略下：

$u_d = u(C, C, C)$

$= [P \cdot 0 + (1-P) \cdot 0] + [P(-8) + (1-P)(-8)] + [P \cdot (-8) + (1-P)(-8)]$

$= -16$

	t=1	t=2	t=3
非理性囚徒 1	D	C	C
理性囚徒 1	D	C	C
囚徒 2	C	C	C

要求$u_a \geq u_b \leftrightarrow 17P - 19 \geq 8P - 17 \leftrightarrow P \geq 2/9$

要求$u_a \geq u_c \leftrightarrow 17P - 19 \geq 8P - 18 \leftrightarrow P \geq 1/9$

要求$u_a \geq u_d \leftrightarrow 17P - 19 \geq -16 \leftrightarrow P \geq 3/17$

因$\frac{3}{17} < \frac{2}{9}$

故当$P \geq 2/9$时，$u_a \geq \begin{cases} u_b \\ u_c \\ u_d \end{cases}$

故a为最优战略。

其次，再证明给定囚徒 2 的战略，理性囚徒 1 的战略为最优战略。

对于理性囚徒 1，当$t = 1$时，他若选“C”，则在$t = 2$时，囚徒 2 知道了其理性人属性，通过贝叶斯法则修正先验概率，得后验概率$\tilde{P} = 0$。

	t=1	t=2	t=3
理性囚徒 1	D	C	C
囚徒 2	D	D	C

$t = 2$时囚徒 2 将选“C”而非“D”。

理性囚徒 1 在$t = 1$选“C”是非均衡路径上的。

均衡要求理性囚徒 1 在$t = 1$选“D”。

当$t=2$时，因$t=3$的博弈结果是确定的，故理性囚徒 1 选“C”是占优的，故选“C”。

故理性囚徒 1 的战略也是最优的（给定囚徒 2 的战略及后验概率）。

事实上，当$t=1$理性囚徒 1 选“C”时，其支付为$0+(-8)+(-8)=-16$，而选“D”时的支付为$(-1)+0+(-8)=-9>-16$，故选“D”为最优。

$T=3$时定理 1 成立。

归纳法假设：

设当$T\geq 3$时定理 1 成立，下面证$T=T+1$时定理 1 也成立。

首先证明这样一个结论，即从$t=T-2$开始的博弈中，囚徒 2 的最优战略仍是（D,D,C）。显然，当非理性囚徒 1 在$t=T-2$的选择为 D 时（对应于囚徒 2 在$t=T-3$的选择为 D），这个博弈与前面讨论的$T=3$时的博弈相同，故已证（D,D,C）为囚徒 2 的最优战略。

故下面只需证非理性囚徒 1 在$t=T-2$的选择为 C（对应于囚徒 2 在$t=T-3$的选择 C）时，（D,D,C）仍为囚徒 2 的最优战略。下面计算囚徒 2 在各种可能战略选择下的支付（只是$t=T-3$后博弈的支付，不是整个博弈的支付）。

$$
\begin{aligned}
u(C,C,C)&=[P(-8)+(1-P)\cdot 0]+[P\cdot(-8)+(1-P)(-8)]\\
&\quad+[P(-8)+(1-P)(-8)]\\
&=-8P-16
\end{aligned}
$$

	t=1	t=2	t=3
非理性囚徒 1	C	C	C
理性囚徒 1	D	C	C
囚徒 2	C	C	C

$$
\begin{aligned}
u(C,D,C)&=[P(-8)+(1-P)\cdot 0]+[P\cdot(-10)+(1-P)(-10)]\\
&\quad+[P\cdot 0+(1-P)(-8)]\\
&=-18
\end{aligned}
$$

	t=1	t=2	t=3
非理性囚徒 1	C	C	C
理性囚徒 1	D	C	C
囚徒 2	C	D	C

$$
\begin{aligned}
u(D,C,C)&=[P(-10)+(1-P)(-1)]+[P\cdot 0+(1-P)(-8)]\\
&\quad+[P\cdot(-8)+(1-P)(-8)]\\
&=-P-17
\end{aligned}
$$

	t=1	t=2	t=3
非理性囚徒 1	C	D	C
理性囚徒 1	D	C	C
囚徒 2	D	C	C

$$u(D,D,C)=[P(-10)+(1-P)(-1)]+[P\cdot(-1)+(1-P)(-10)]$$
$$+[P\cdot 0+(1-P)(-8)]$$
$$=8P-19$$

	t=1	t=2	t=3
非理性囚徒 1	C	D	D
理性囚徒 1	D	C	C
囚徒 2	D	D	C

因为有：$8P-19\geq -8P-16 \leftrightarrow P\geq \frac{3}{16}$

因 $\frac{3}{17}<\frac{2}{9}$，所以由 $P\geq\frac{2}{9}>\frac{3}{16}$，故有

$u(D,D,C)>U(C,C,C)$

又 $8P-19\geq 18\leftrightarrow P\geq\frac{1}{8}$

因 $\frac{1}{8}<\frac{2}{9}\leq P$

$u(D,D,C)>u(C,D,C)$

又 $8P-19\geq -P-17\leftrightarrow P\geq\frac{2}{9}$

故 $u(D,D,C)>u(D,C,C)$

所以 (D,D,C) 为最优战略

这说明，给定理性囚徒的上述战略，囚徒 2 的最优战略在 $t=T-2$ 后必为 (D,D,C)。

进一步，上面还证明了这样一个结论，即当 $T=3$ 时，给定理性囚徒 1 的战略为 (D,C,C)，无论 $t=T-2\ (=1)$ 时非理性囚徒 1 的选择是 C 或是 D，囚徒 2 的最优战略必为 (D,D,C)。

下面证明这一结论可推广到 $T\geq 3$，用数学归纳法，已证 $T=3$ 时成立，设 T 时成立，则在 $T=T+1$ 时。

	t=1	t=2	t=3	……	t=T-2	t=T-1	t=T
非理性囚徒 1	C	1.C	D	……	D	D	D
		2.D	D	……	D	D	D
理性囚徒 1	D	D	D	……	D	C	C
囚徒 2	1.C	1.D	D	……	D	D	C
	2.D	2.D	D	……	D	D	C

	t=1	t=2	t=3	……	t=T-2	t=T-1	t=T
非理性囚徒 1	D	1.C	D	……	D	D	D

		2.D	D	……	D	D	D
理性囚徒 1	D	D	D	……	D	C	C
囚徒 2	1.C	1.D	D	……	D	D	C
	2.D	2.D	D	……	D	D	C

若$t=1$时非理性囚徒 1 的选择为 C，此时对应于囚徒 2 的 2 种可能选择 C、D，我们分别用 1 和 2 标明。

由归纳法假设，当$t \ge 2$后，囚徒 2 的任何战略选择都没有选（D,D,⋯,D,C）带来的支付大。

而对应于 1 和 2，囚徒 2 和非理性囚徒 1 在$t \ge 3$后的选择都是一样的。

故只须比较$t=1,2$时囚徒 2 的不同选择的支付差别就可判别孰优孰劣。

当 1 时，囚徒 2 在$t=1,2$的阶段支付和为

$[P(-8)+(1-P)\cdot 0]+[P\cdot(-10)+(1-P)(-1)]=-17P-1$

当 2 时，囚徒 2 在$t=1,2$的阶段支付和为

$[P(-10)+(1-P)(-1)]+[P\cdot(-10)+(1-P)(-1)]=11P-2$

欲使$11P-2 \ge -17P-1 \leftrightarrow P \ge \frac{1}{28}$

因$P \ge \frac{2}{9} > \frac{1}{28}$

故 2 是囚徒 2 的最优选择。

同样，当非理性囚徒 1 在$t=1$选 D 时，对应于囚徒 2 在$t=1$的 2 种选择，他选 D 时在$t=1,2$的阶段支付和为-2，而选 C 时在$t=1,2$的阶段支付和为$-9P-1$。

欲使$-2 > -9P-1 \leftrightarrow P > \frac{1}{9}$

因$P \ge \frac{2}{9} > \frac{1}{9}$，故选 D 是最优的。

所以，无论非理性囚徒 1 在$t=1$选什么，囚徒 2 选（D,D,⋯,D,C）是囚徒 2 的最优战略。

由归纳法假设知对任意$T \ge 3$，无论非理性囚徒 1 在$t=1$选什么，（D,D,⋯,D,C）总是囚徒 2 的最优战略。

下面，我们来完成前一个数学归纳法的证明。给定非理性人 1 在$t=1$选 D，囚徒 2 选（D,D,⋯,D,C）显然是最优战略，因为已证此时非理性人 1 在$t=1$选 D 和 C 时，囚徒 2 选（D,D,⋯,D,C）都是最优战略，当然在非理性人在$t=1$选 D 时仍为最优。

最后，我们来证明给定囚徒 2 的战略，理性囚徒 1 的战略为最优战略。

显然，给定囚徒 2 的战略，理性囚徒 1 若在$t=1$选 C，则支付为$0+T(-8)=-8T$，若选 D，在$t=1$阶段支付为（-1），在$t \ge 2$后，已证$T=3$时（D,D,⋯,D,C,C）为最优，

故建立归纳法假设 $t \geq 2$ 后的最优选择为（D,D,…,D,C,C），则此时支付为 $(-1)+(T-2)(-1)+0+(-8)=-T-7$

欲使 $-T-7 \geq -8T \leftrightarrow T \geq 1$（显然成立）

定理证毕！

由定理 7.1 看出，只要 $T>3(P \geq \frac{2}{9}$ 时)，非合作阶段的总数量总等于 2，与 T 无关。

另外，$P \geq \frac{2}{9}$ 既是有限次囚徒困境不完全信息重复博弈存在合作精炼贝叶斯均衡的充分条件（上已证明），也是必要条件（不难证明）。

非理性囚徒 1				C(D)	C(D)
理性囚徒 1	D	……	D(C)	C	C
囚徒 2			C(D)	C(D)	C

事实上，$P<\frac{2}{9}$，在 $t=T-1$，无论非理性囚徒 1 选什么（即无论囚徒 2 在 $t=T-2$ 选什么），囚徒 2 在 $t=T-1$ 的最优选择总是 C。

因为，当非理性囚徒 1 在 $t=T-1$ 选 C 时，囚徒 2 在 $t=T-1$ 选 C 时所带来的 $t=T-1$ 及之后阶段的支付和为 $(-8)+(-8)=-16$

此时囚徒 2 在 $t=T-1$ 选 D 时所带来的 $t=T-1$ 及之后阶段的支付和为：

$[P(-10)+(1-P)(-10)]+[P \cdot 0+(1-P)(-8)]=8P-18$

欲使 $-16>8P-18 \leftrightarrow P<\frac{2}{8}$

因 $P<\frac{2}{9}<\frac{2}{8}$

故此时囚徒 2 在 $t=T-1$ 选 C 为最优。

当非理性囚徒 1 在 $t=T-1$ 选 D 时，囚徒 2 在 $t=T-1$ 选 C 时带来的 $t=T-1$ 及之后的阶段支付和为：

$[P \cdot 0+(1-P)(-8)]+(-8)=8P-16$

此时囚徒 2 在 $t=T-1$ 选 D 时带来的 $t=T-1$ 及之后的阶段支付和为：

$[P(-1)+(1-P)(-10)]+[P \cdot 0+(1-P)(-8)]=17P-8$

欲使 $8P-16>17P-18 \leftrightarrow P<\frac{2}{9}$，当然成立。

无论非理性囚徒 1 在 $t=T-1$ 选什么（即无论囚徒 2 在 $t=T-2$ 选什么），囚徒 2 在 $t=T-1$ 选 C 为最优。

既然囚徒 2 在 $t=T-1$ 总会选 C，理性囚徒 1 就没有必要在 $t=T-2$ 选 D 了，故他会选 C。

从$t = T - 2$开始的博弈等价于从$t = T - 1$开始的阶段博弈的二阶段重复博弈。

囚徒 2 在$t = T - 2$总会选 C。

类似反推直至$t = 1$，囚徒 2 总选 C。

囚徒 1（无论是否理性）在各阶段也总选 C。

唯一的子博弈精炼贝叶斯均衡是两个囚徒（除了非理性囚徒 1 在第一阶段选 D(实际不出现，因囚徒 1 实际上是理性的)）在各阶段都选 C 的不合作均衡。

在上面，我们只是假定其中一个囚徒的类型是私人信息，此时当 P 充分大之后才存在合作解。但是，下面将看到，当假定两个囚徒的类型都是私人信息时，只要 P>0 就可保证合作均衡的存在（只要重复博弈的次数足够多）。

下面对此进行一个大致的说明：

这里假定非理性囚徒 1 的战略是“冷酷战略”，即开始选 D（合作），直到对方在t阶段选 C（不合作），然后从$t + 1$开始选 C 直到 T（“针锋相对”战略代表的是“你不仁我不义”，“投之以桃，报之以李”，而冷酷战略代表的是绝不原谅对方的任何背信弃义行为，这里用“冷酷战略”取代“针锋相对”战略来描绘非理性囚徒 1 的战略，只是为了简化分析，与结论无关）。

我们来看理性囚徒 1，他在开始阶段有两个可能的选择，即选 C 或 D，先看他选 C 的情形，此时因暴露了类型（因非理性囚徒 1 开始是选 D），理性囚徒 2 将在$t = 2$及之后各阶段都选 C，而根据冷酷战略，非理性囚徒 2 也在$t = 2$及之后各阶段都选 C。

给定这种情形，理性囚徒 1 在$t = 2$及之后各阶段的最优选择也是选 C。

显然，这是$t = 2$及之后各阶段博弈的唯一均衡，故此时理性囚徒 1 的最大期望支付为：

$u_1(C) = 0$（当囚徒 2 选 D）$+ (-8)(T - 1) = -8(T - 1)$

再看理性囚徒 1 开始选 D 的情形：在他开始选 D 的情形，可以有多种可选择的战略，我们为了证明他不会在开始选 C，只须找出一个开始选 D 的战略，其支付大于他开始选 C 所带来的最大支付即可。

我们只需看看这样一个战略：

开始选 D 直到 T，除非囚徒 2 在某t阶段选 C，此时，在$t + 1$开始选 C 直到 T（即冷酷战略）（这个战略实际上并非囚徒 1 的最优战略，但我们将证明，只要 T 充分大，即使是该非最优战略也要优于开始就选 C 的战略）。

当囚徒 2 是非理性的，囚徒 1 从该战略中获得支付$(-1)T$；当囚徒 2 是理性的时，囚徒 1 从该战略中获得的最小支付为当囚徒 2 在$t = 1$选 C，随后两人都选 C 的支付，即$(-10) + (-8)(T - 1) = -8T - 2$

理性囚徒 1 从该战略得到的最小期望支付为：

$P(-T)+(1-P)(-8T-2)$

当 $P(-T)+(1-P)(-8T-2)>-8(T-1)\leftrightarrow T>(3-2P)/7P$

该冷酷战略优于开始就选 C 的战略。

因 $T^*=(3-2P)/7P=\frac{3}{7P}-\frac{2}{7}$ 是 P 的减函数，故 P 愈大，T^* 就愈小，合作均衡就愈可能出现（因这里的博弈是对称的，两个囚徒在开始都不会选 C，同样的分析适用于 $t=2$，$t=3$，…，$t=T^*-1$，即他们都不会在这些阶段中选 C 而是选 D，即相互合作）。

显然，只要 $P>0$，无论 P 多么小，合作行为在 $T>T^*$ 的博弈中总会出现，只不过 P 很小时，合作行为出现对 T 的要求很大，当 $P\to 0$ 时，$T^*\to\infty$。

一般地，我们有下述定理：

定理 7.2 （KMRW 定理）

在 T 阶段重复囚徒困境博弈中，如果每个囚徒都有 $P>0$ 的概率是非理性的（即对方认为其只选择“针锋相对”或“冷酷战略”），如果 T 足够大，那么存在一个 $T_0<T$，使得下列战略组合构成一个精炼贝叶斯均衡：所有理性囚徒在 $t\le T_0$ 的各阶段选 D（合作），在 $t>T_0$ 各阶段选 C（不合作）；并且非合作阶段的数量 $(T-T_0)$ 只与 P 有关而与 T 无关。

该定理这里不给出证明，证明请参见 kreps, Milgrom, Roberts 及 Wilson（1982）。

KMRW 定理给出了对声誉的解释。声誉的作用是诱使对方今后与之合作，但当 $t\ge T_0$ 后，声誉的使用用尽，没有必要付出短期利益去维护声誉，图穷匕首见，出现不合作行为。

建立和维护声誉是有成本的，并且只有充分关注长期利益者才会去建立和维护声誉。

因为名牌产品对企业来说是一种声誉，因建立和维护声誉需要付出成本，故名牌产品的价格一般要比非名牌产品的价格高一些（即使质量是一样的）。

KMRW 定理为现实中诸多合作现象提供了一个强有力的解释。如寡头市场上的合谋行为（在完全信息下是不可能出现的）、第五章中的“连锁店悖论”等。另外，有意思的是 KMRW 定理为我们中国人推崇的“大智若愚”行为提供了合理解释。“智”者，理性人也，为了诱使对方合作，智者可以表现得象个愚者（非理性）那样。

注意，P=0 时只有不合作解，P>0 就在 T 充分大时有了合作解，这是信息经济学中参数在临界点处变化带来行为不连续性诸多例子中的一个。

7.5.2 声誉模型的应用：货币政策

我们在第五章中完全信息动态博弈模型中看到，货币政策存在动态不一致性，即公众不会相信政府的零通胀承诺，从而使政府陷入通胀与失业的双重困境中。但 Barro（1986）和 Vickers（1986）运用 KMRW 声誉模型证明：如果公众有关政府偏好的信息是不完全的，出于声誉方面的考虑，政府可能选择不制造通胀，即使政府的任期是有限的（有限重复博弈）。我们在第 7 章中的 7.3.3 中已对此加以了一种介绍，这里再按上一节的方式作一说明。

假定

公众与政府构成博弈中的两个局中人，而公众认为政府可能有两种类型，即强政府和弱政府。

强政府——从不制造通胀

弱政府——在一定压力下（如利益集团的压力）可能会通过制造通胀来解决短期矛盾。

但弱政府可通过假装强政府来建立一个不制造通胀的声誉。

公众不知道政府的类型，但可通过观测通胀率来推断政府的类型。

一旦政府制造了通胀，公众就认为政府是弱政府，从而在公众对通胀率的理性预期（准确预知）下，政府在随后阶段的通胀不能带来任何产出或就业的增加。

以下考察在什么情况下，弱政府不选择制造通胀（真实的政府类型是弱政府，但公众以正的概率认为有可能出现强政府，合作均衡意味着弱政府也会选择不制造通胀）。

假定

政府的单阶段效用函数如下：

$$u = -\frac{1}{2}\pi^2 + b(\pi - \pi^e)$$

其中 π 为实际通胀率（政策变量），π^e 是公众预期的通胀率，b 为政府的类型。

$$b\begin{cases} = 0 & \text{代表强政府} \\ = 1 & \text{代表弱政府} \end{cases}$$

假定　公众对政府为强政府（b=0）的先验概率为 P。

认为是弱政府（b=1）的先验概率为（$1 - P_0$）

弱政府在单阶段博弈中的最优通胀率计算如下：

$$\frac{du}{d\pi} = -\pi + b = 0$$

得 $\pi = b = 1$

最大效用 $u = -\frac{1}{2}\times 1^2 + 1\times(1-1) = -\frac{1}{2}$（理性预期下）

假定

博弈重复 T 阶段，设 y_t 为 t 阶段弱政府选择零通胀的概率（即假装强政府的概率），x_t 为公众认为弱政府选择零通胀的概率；在均衡中，$x_t = y_t$（公众和政府都知道均衡，且均衡是共同知识，当然公众知道 y_t，故 $x_t = y_t$，即公众知道均衡下的 y_t，故有均衡下的 $x_t = y_t$）。

如果公众在 t 阶段未预测到通胀，则由贝叶斯法则，公众在 $t+1$ 阶段认为政府是强政府的后验概率为：

$$\tilde{P}_{t+1}\left(b=0 \middle| \pi_t = 0\right) = \frac{\Pr ob(b=0, \pi_t = 0)}{\Pr ob(\pi_t = 0)}$$

$$= \frac{\Pr ob(\pi_t = 0 \mid b=0)\Pr ob(b=0)}{\Pr ob(\pi_t = 0 \mid b=0)\Pr ob(b=0) + \Pr ob(\pi_t = 0 \mid b=1)\Pr ob(b=1)}$$

因

$$= \frac{1\times \tilde{P}_t}{1\times \tilde{P}_t + x_t(1-\tilde{P}_t)}$$

$1\times \tilde{P}_t + x_t(1-\tilde{P}_t) \in [x_t, 1]$

故 $1\times \tilde{P}_t + x_t(1-\tilde{P}_t) \le 1$

$\tilde{P}_{t+1}\left(b=0 \middle| \pi_t = 0\right) \ge \tilde{P}_t$

意即：若政府未在 t 阶段制造通胀，则公众在 $t+1$ 阶段认为政府是强政府的概率向上调整。

如果政府在 t 阶段制造了通胀，则

$$\tilde{P}_{t+1}\left(b=0 \middle| \pi_t = 0\right) = \frac{\Pr ob(\pi_t = 1 \mid b=0)\Pr ob(b=0)}{\Pr ob(\pi_t = 1 \mid b=0)\Pr ob(b=0) + \Pr ob(\pi_t = 1 \mid b=1)\Pr ob(b=1)} = 0$$

即一旦观测到了通胀，公众就认为政府是弱的。

下面，我们运用逆向归纳法的思路来导出均衡解。

先看 $T-1$ 和 T 阶段的解。

在最后的 T 阶段，维护不制造通胀的声誉已无必要。

弱政府此时的选择必为 $\pi = b = 1$，公众预期通胀率 $\pi_T^e = 1(1-\tilde{P}_T) = 1-\tilde{P}_T$

弱政府的阶段效用水平为：

$$u_T = -\frac{1}{2}\pi_T^2 + (\pi_T - \pi_T^e)$$

$$= -\frac{1}{2} + [1-(1-\tilde{P}_T)] = \tilde{P}_T - \frac{1}{2}$$

u_T 是 $\tilde{P}_T$ 的增函数，即弱政府在最后阶段的效用是其声誉的增函数

这就是为什么弱政府有积极性建立强政府声誉的原因。

再看 $T-1$ 阶段。假定弱政府在 $T-1$ 之前未制造通胀

$\tilde{P}_{T-1} > 0$

公众的通胀率预期为

$\pi^e_{T-1} = 1 \times (1-\tilde{P}_{T-1})(1-x_{T-1})$

其中：1 是弱政府的最优通胀率，$1-\tilde{P}_{T-1}$是政府是弱政府的概率，$(1-x_{T-1})$是公众认为弱政府选择通胀的概率。

令δ为政府的贴现因子。为简单计，先考虑纯战略博弈，即弱政府或选择零通胀或选择通胀的情形

$y_t = 0$或 1 的情形。（我们知道，在混合战略博弈均衡中，对应有带来相同期望效用的多个最优纯战略，故知道了最优纯战略的条件也就知道了混合战略的条件）。

如果弱政府在$T-1$选择通胀（$y_{T-1}=0$，即$\pi_{T-1}=1$）

则$\tilde{P}_T = 0$（T 阶段公众知道政府是弱的）。

给定公众的通胀预期π^e_{T-1}，弱政府此举的二阶段总效用为：

$$u_{T-1}(1)+\delta u_T(1) = \left[-\frac{1}{2}+(1-\pi^e_{T-1})\right]-\frac{1}{2}\delta = \frac{1}{2}-\pi^e_{T-1}-\frac{1}{2}\delta$$

如果弱政府在$T-1$选不制造通胀（$y_{T-1}=1$，即$\pi_{T-1}=0$），则此举的二阶段总效用为

$$u_{T-1}(0)+\delta u_T(1) = -\pi^e_{T-1}+\delta\left(\tilde{P}_T-\frac{1}{2}\right)$$

故当且仅当下列条件成立时，选$\pi_{T-1}=0$要优于$\pi_{T-1}=1$：

$$-\pi^e_{T-1}+\delta\left(\tilde{P}_T-\frac{1}{2}\right) \geq \frac{1}{2}-\pi^e_{T-1}-\frac{1}{2}\delta \leftrightarrow \tilde{P}_T \geq \frac{1}{2\delta}$$

因在均衡下，有$x_{T-1}=y_{T-1}$

若$y_{T-1}=1$构成弱政府的均衡战略，则$x_{T-1}=y_{T-1}=1$

$P_T = P_{T-1}$（公众未能获得更多的信息）

上述条件意味着有

$$\tilde{P}_{T-1} \geq \frac{1}{2\delta}$$

意即：如果公众在$T-1$阶段认为政府是强政府的概率不小于$\frac{1}{2\delta}$，则弱政府就会假装强政府。

即政府的声誉越好，则政府维持声誉的积极性就愈高。

若$\tilde{P}_T = \frac{1}{2\delta}$，则上述分析表明了$T-1$阶段弱政府假装或不假装强政府都是最优的。

它以任意概率分布随机选$y_{T-1}=0$或 1 都是最优的混合战略

任何$y_{T-1}\in[0,1]$都是最优的。

因均衡要求$x_{T-1}=y_{T-1}$，将$\tilde{P}_T = \frac{1}{2\delta}$代入前述贝叶斯法则公式。

$$\tilde{P}_T = \tilde{P}_T\left(b=0 \middle| \pi_{T-1}=0\right) = \frac{\tilde{P}_{T-1}\times 1}{\tilde{P}_{T-1}\times 1+(1-\tilde{P}_{T-1})x_{T-1}}$$

$$\frac{1}{2\delta}=\frac{\tilde{P}_{T-1}}{\tilde{P}_{T-1}+(1-\tilde{P}_{T-1})x_{T-1}}$$

$$y_{T-1}=x_{T-1}=\frac{(2\delta-1)\tilde{P}_{T-1}}{1-\tilde{P}_{T-1}}$$

因 x_{T-1} 是 $\tilde{P}_{T-1}$ 的增函数，公众愈是认为政府是强政府，弱政府选择不制造通货膨胀的概率就愈高（假定 $\delta>\frac{1}{2}$，即 $2\delta-1>0$）。

当 $\tilde{P}_{T-1}\to\frac{1}{2\delta}$ 时， $y_{T-1}\to 1$

下面用数学归纳法证明：如果政府在 $T-1$ 阶段选零通胀是最优的（之前各阶段未制造通胀），则在所有 $t<T-1$ 阶段选零通胀都是最优的（之前各阶段未制造通胀）。

已证当 $\tilde{P}_T\geq\frac{1}{2\delta}$ 时，政府在 $T-1$ 阶段选零通胀最优的（之前各阶段未制造通胀）。

归纳法假设

设 $t\leq T-1$ 时，政府在 t 阶段选零通胀是最优的（之前各阶段未制造通胀），则在 $t=T-1$ 时，若政府选通胀，则自 $t-1$ 及之后的各阶段效用贴现和为（$t-1$ 之后因类型暴露，只有不断制造通胀）：

$$\begin{aligned}S_{t-1}(t)&=u_{t-1}(1)+\delta u_t(1)+\cdots+\delta^{T-t+1}u_T(1)\\&=-\frac{1}{2}(1-\pi_{t-1}^e)-\frac{1}{2}\delta-\cdots-\frac{1}{2}\delta^{T-t+1}\\&=\frac{1}{2}-\pi_{t-1}^e-\frac{1}{2}\delta-\cdots-\frac{1}{2}\delta^{T-t+1}\end{aligned}$$

若此阶段政府不制造通胀，则自 $t-1$ 及之后的各阶段效用贴现和为：

$$S_{t-1}(0)=u_{t-1}(0)+\delta S_t(0)$$

该式成立的理由是，由归纳法假设知在 t 阶段，当之前各阶段都未制造通胀时，政府选不制造通胀是最优的，故当政府在 $t-1$ 阶段未选通胀时，必在 t 阶段也会选零通胀。下面证明必有

$$S_{t-1}(0)\geq S_{t-1}(1)\leftrightarrow u_{t-1}(0)+\delta S_t(0)\geq\frac{1}{2}-\pi_{t-1}^e-\frac{1}{2}\delta-\cdots-\frac{1}{2}\delta^{T-t+1}$$

由归纳法假设有：

$$\begin{aligned}S_t(0)&\geq u_t(1)+\delta u_{t+1}(1)+\cdots+\delta^{T-t}u_T(0)\\&=-\frac{1}{2}+(1-\pi_t^e)-\frac{1}{2}\delta-\cdots-\frac{1}{2}\delta^{T-1}\\&=\frac{1}{2}-\pi_t^e-\frac{1}{2}\delta-\cdots-\frac{1}{2}\delta^{T-t}\end{aligned}$$

故

$$S_{t-1}(0)=u_{t-1}(0)+\delta S_t(0)$$
$$=-\pi_{t-1}^e+\delta S_t(0)$$
$$\geq -\pi_{t-1}^e+\delta\left(\frac{1}{2}-\pi_t^e-\frac{1}{2}\delta-\cdots-\frac{1}{2}\delta^{T-t}\right)$$

欲证 $S_{t-1}(0)\geq S_{t-1}(1)$

只需证 $-\pi_{t-1}^e+\delta\left(\frac{1}{2}-\pi_t^e-\frac{1}{2}\delta-\cdots-\frac{1}{2}\delta^{T-t}\right)\geq\frac{1}{2}-\pi_{t-1}^e-\frac{1}{2}\delta-\cdots-\frac{1}{2}\delta^{T-t+1}$

$$\leftrightarrow \delta\left(\frac{1}{2}-\pi_t^e\right)\geq\frac{1}{2}-\frac{1}{2}\delta$$

因 $-\pi_t^e=-(1-\tilde{P}_t)(1-x_t)\geq-(1-\tilde{P}_t)=\tilde{P}_t-1$（因 $x_t\in[0,1]$）

故 $\delta\left(\frac{1}{2}-\pi_t^e\right)\geq\delta\left(-\frac{1}{2}+\tilde{P}_t\right)=\delta\tilde{P}_t-\frac{1}{2}\delta$

前式成立 $\rightarrow\delta\tilde{P}_t\geq\frac{1}{2}\leftrightarrow\tilde{P}_t\geq\frac{1}{2\delta}$

由命题在 $t=T-1$ 成立知已有 $\tilde{P}_t\geq\frac{1}{2\delta}$

由归纳法假设，知政府在 $t\geq t$ 后各阶段都选零通胀

$x_t=1(t\geq t)$

由前述贝叶斯法则

$$\tilde{P}_{t+1}=\tilde{P}_{t+1}(b=0\mid\pi_t=0)=\frac{\tilde{P}_t\times1}{\tilde{P}_t\times1+(1-\tilde{P}_t)x_t}=\tilde{P}_t$$

$$\tilde{P}_t=\tilde{P}_{t+1}=\cdots=\tilde{P}_T\geq\frac{1}{2\delta}$$

由此知上述条件成立

证毕！

我们于是有下述均衡：

若 $P_0\geq\frac{1}{2\delta}$，则存在如下精炼贝叶斯均衡（是混同均衡）：

强政府选择：$\pi_0=\pi_1=\cdots=\pi_{T-1}=\pi_T=0$；

弱政府选择：$\pi_0=\pi_1=\cdots=\pi_{T-1}=0,\pi_T=1$；

公众预期的通胀为：

$\pi_0^e=\pi_1^e=\cdots=\pi_{T-1}^e=0$；$\pi_T^e=1-\tilde{P}_T=1-P_0$；

后验概率：

在均衡路径上，$\tilde{P}_T=\tilde{P}_{T-1}=\tilde{P}_{T-2}=\cdots=P_0$；

在非均衡路径上，若 $\pi_1=1$ 则 $P_{t+1}=0$（阻止政府选择通胀）。

政府的支付为 $\begin{cases} 强政府:\sum_{t=1}^{T}\delta^t u_t = 0+0+\cdots+0=0 \\ 弱政府:\sum_{t=1}^{T}\delta^t u_t = 0+\cdots+0+\delta^T\left(P_0-\frac{1}{2}\right)=\delta^T\left(P_0-\frac{1}{2}\right) \end{cases}$

上述讨论的一个基本结论是：信息结构的微小变化对政府的行为有重要的影响。

7.6.效率工资理论

经典的经济学理论认为工资率水平由工人的边际产值决定，但观察到的一些事实发现很多企业给予工人的工资率远高于其边际产值。同时，经济学家观察到日本企业在员工流动频率上远低于西方国家。效率工资理论试图对这些观察事实给予解释。

效率工资的含义是指高工资带来高产出，高工资是激励工人努力工作的手段。

首先，用高工资可以吸引高素质的人才，因为高素质人才的市场评价高，只有足够高的工资才会招聘到高素质人才。高素质人才就会带来高产出。

其次，高工资有激励作用，因为工人知道，不努力工作，不尽量做到最好，其高工资职位就会被他人取代。所以，高工资可以激励努力工作。在发展中国家，高工资收入还可为工人提供更好的营养，从而带来高产出；在发达国家，高工资收入可吸引更多有能力的工人到企业求职，或可激励现有工人更加努力工作。

我们下面介绍由 Shapiro 和 Stiglitz 提出的效率工资模型。

Shapiro-stiglitz 模型（1984）：这是一个无限次重复博弈模型。

局中人 1——企业

局中人 2——工人

局中人 N——自然

博弈的大意是：企业为激励工人努力工作，一方面支付高工资，同时又威胁工人：一旦发现其偷懒，就立即开除他。

因为企业支付高工资，故企业减少了对劳动力的需求，造成部分工人的高薪就业，但其它工人成为非自愿企业。因为即使工人只要求低工资，仍然不能被雇用，因企业实行高工资低就业策略。

这样，失业工人数量愈多，一个被解雇的工人寻找新工作所需时间就愈长，故而解雇的威胁就更加有效。

在竞争均衡条件下，工资水平 w 和失业率 u 恰好可以使工人不敢偷懒，并且企业在工资水平 w 时的劳动需求恰好使失业率等于 u 。

其过程是：在工资上调过程中，由于工人怕因偷懒而被解雇失去高薪，同时又因工

资上升减少了对工人的需求导致失业率上升，找工作代价大而使其保留支付下降，这种双层作用激励工人更加努力工作，在均衡时，劳动需求使失业率为u，且和w使工人不敢偷懒。

下面，我们只分析模型的一部分，即一个企业与一个工人间的重复博弈，而不考虑竞争均衡。

1．阶段博弈G：行动顺序如下：

①企业开出一个工资水平w；

②工人选接受或拒绝，若拒绝，则工人成为自我雇佣者，工资水平为w_0（保留支付）。若接受，则工人选择是努力工作（会带来e单位负效用）还是偷懒（不带来任何负效用）。

企业不能直接预测到工人的努力程度（或工人在选择了接受后的决策），但企业和工人都可预测到工人的产出水平。

假定：只有二个产出水平：$\begin{cases} 高产出 y>0 \\ 低产出 y=0 \end{cases}$

当工人努力工作时，必为高产出y。

当工人偷懒时，则产出为高产出y的概率为p。

产出为低产出$y=0$的概率为$(1-p)$（因为影响产出的因素除工人努力程度外，还有其它不可控的随机因素，如气候等，正是如此，工人努力程度才是不可观测的，否则，若产出与工人努力程度 1-1 对应，则可直接经观测产出水平来观测工人努力程度）。

显然，若预测到低产出，则直接可知工人在偷懒（低产出是偷懒无可辩驳的证据）。

倘若企业以w的工资雇佣了工人，则当工人努力工作时，产出为高产出，局中人支付为：

企业支付$=y-w$

工人支付$=w-e$

当工人偷懒时，$\begin{cases} 若产出为低产出 y=0,\ 企业支付=\text{-}w,\ 工人支付=w \\ 若产出为高产出 y>0,\ 企业支付=y\text{-}w,\ 工人支付=w \end{cases}$

假定有$y-e>w_0>py$（$=py+(1-p)\cdot 0$，即工人偷懒时的期望产出水平），其意为：对工人来说，在企业努力工作比自我雇佣（失业）更好（当企业付给工人足够高工资且工资不超过工人的边际产值y时，工人受雇于企业比失业更好），此即$y-e>w_0$的意义；而$w_0>py$的意义是工人自我雇佣要比在企业中工作并偷懒且被企业知道其偷懒而最多支付其$py+(1-p)\cdot 0=py$的工资为好。

阶段博弈树：

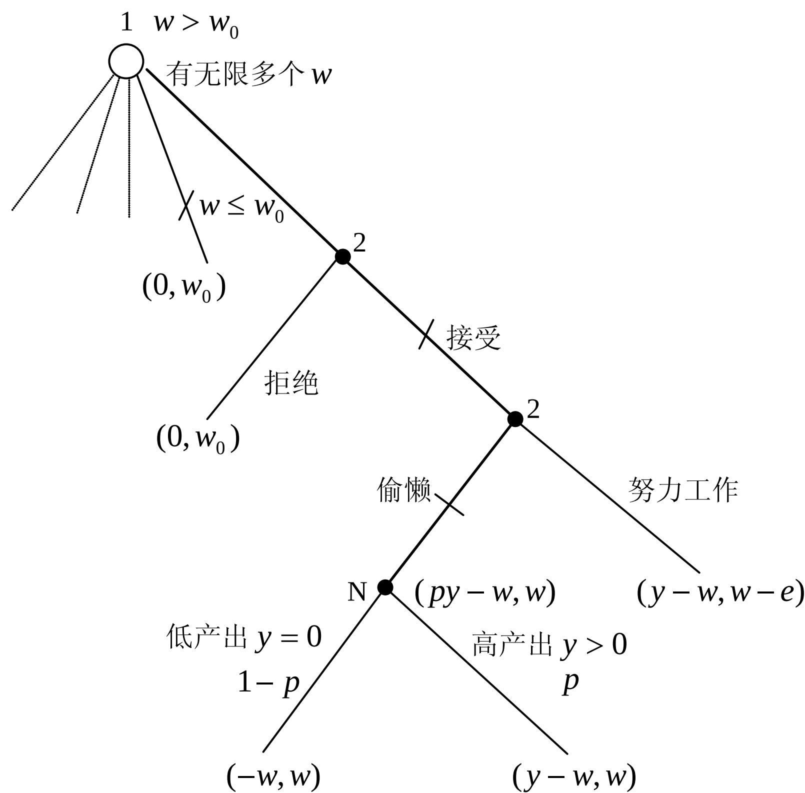


图 7.14 效率工资模型阶段博弈树

阶段博弈的子博弈精炼均衡

因为有假设 $w_0 > py \Leftrightarrow py - w_0 < 0$，又因在 $w > w_0$ 时有 $py - w < py - w_0 < 0$，故由逆向归纳法知阶段博弈的子博弈精炼均衡结果必为：

$$\begin{cases} 企业选 w \le w_0 (可选 w = 0) \\ 工人选自我雇佣 \end{cases}$$

支付 $= (0, w_0)$

精炼均衡由图 7.14 表示。

精炼均衡是：因为企业知道它给出 $w > w_0$ 时工人也会偷懒，此时工人没有动机和积极性去努力工作（事后机会主义、道德风险、短期行为），故企业在开始就会开出 $w \le w_0$（$w = 0$），工人选自我雇佣。

2．无限重复博弈

若上述阶段博弈无限次重复，则合作结果就可能出现。此时，企业开出足够高的 w 会激励工人努力工作，因为此时工人会考虑长远利益，事后机会主义就可能消失。这是企业用合作好处诱使工人合作，否则会因短期利益损失潜在的长远利益。

定义 7.5　当某一决策结之前的产出都是高产出且之前所有的工资开价都恒为某 w^* 且

都被工人接受了，称该决策结之前的博弈过程为“高工资 、高产出”。

企业战略：在第一阶段开出工资率水平 $w = w^*$，在其后的每一阶段，如果博弈过程为“高工资、高产出”，则继续开出工资水平 w^*；否则，开出 $w = 0$（冷酷战略！）

工人战略：若 $w \geq w_0$，则接受企业给出的工资，否则，选拔自我雇佣；当接受企业工资时，若博弈过程（包括本阶段工资）是高工资、高产出，则选努力工作，否则选偷懒。

显然，企业战略本质上就是前述“冷酷战略”；若在前面所有阶段的博弈中都相互合作，就继续合作；但一旦有一次合作被打破，就永远转向阶段博弈的子博弈精炼解。

工人的战略虽也类似于冷酷战略，但由于工人在序贯行动的阶段博弈中的行动在后，其战略也更加灵活一些。在一个基于同时行动阶段博弈的重复博弈中，一方的背离只能在某一阶段结束时才可观测到；不过当阶段博弈是序贯行动时（这里正是如此），首先行动方的背离在同一阶段就可被观测到（并且可对其作出反应）。

故工人的战略将会是：

若前面所有阶段的博弈都合作则继续合作，但若企业一旦有背离就选择自己本阶段的最优行动，因他知道将来所有阶段的博弈都将出现阶段博弈的子博弈精炼解。

即若 $w \neq w^*$，但 $w \geq w_0$，则工人将接受企业的工资但选择偷懒。

下面，我们导出上述双方的战略构成子博弈精炼纳什均衡的有关条件。

①首先导出双方战略成为纳什均衡的条件；

②然后证明它们是子博弈精炼解。

在以下的证明中，充分运用了博弈树在结构上的等价性，即有：

等价性假定

在博弈树的重复的各个不同阶段博弈中，由均衡战略组合给出的从各个阶段博弈开始的子博弈的支付（未来支付贴现和）相等。

假设企业在第一阶段开出的工资 w^* 是均衡工资。

下面看工人在此工资水平下的均衡战略应满足的条件。给定企业的冷酷战略且企业第一阶段开出工资 $w = w^*$，要求工人接受这一工资水平且选择努力工作是最优的，应有什么条件该满足呢？

首先应有 $w^* > w_0$。

若工人选努力工作，则必得高到产出 $y > 0$，进入下一阶段，给定的企业战略给出：企业将再次开出工资水平 w^*，在此后一个阶段博弈中，工人又开始进行与最初阶段博弈（重复的阶段博弈）相同的决策（接受还是拒绝，偷懒还是努力）。

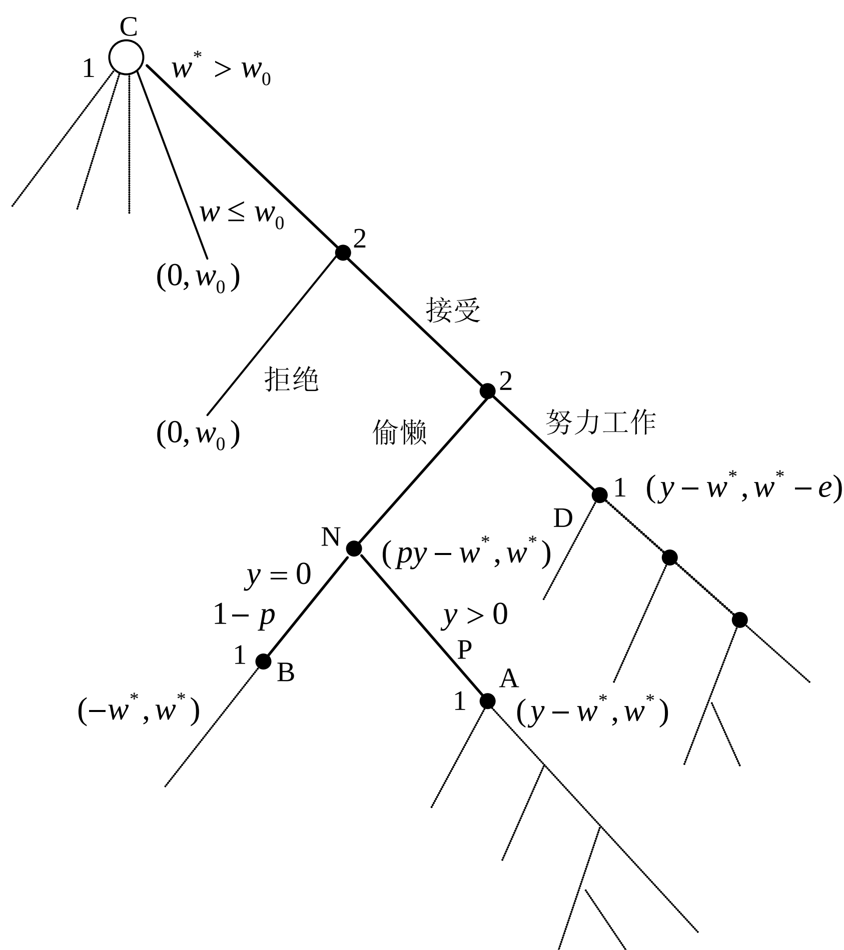


图 7.15 博弈树结构上的等价性

现假定均衡是合作的，即工人总是接受 w^* 且选努力工作，设工人在任一重复阶段博弈中的“子博弈支付现值”为 v_e，则有

$v_e = (w^* - e) + \delta v_e$，其中 δ 为工人的贴现因子

$v_e = (w^* - e)/(1-\delta)$

给定企业战略，倘若工人选偷懒是最优战略，当他选偷懒时，将以 p 的概率带来高产出，到达决策结 A，然后进入下一重复阶段博弈，同时还以 $(1-p)$ 的概率带来低产品，到达决策结 B，之后企业永远开出 $w=0$，工人也将永远选自我雇用，其后工人支付的现值和为 $\dfrac{w_0}{1-\delta}$。

设在这种情形下工人在任一重复阶段博弈中的“子博弈支付现值和”为 v_s，则有：

$$v_s = w^* + \delta\left[pv_s + (1-p)\frac{w_0}{1-\delta}\right]$$

$$v_s = \left[(1-\delta)w^* + \delta(1-p)w_0\right]/(1-\delta p)(1-\delta)$$

在均衡是“合作”的假定下，应有$v_c \ge v_s$，即选择努力工作是最优的，即：

$$w^* \ge w_0 + \frac{1-p\delta}{\delta(1-p)}e = w_0 + \left[1+\frac{1-\delta}{\delta(1-p)}\right]e = (w_0+e) + \frac{1-\delta}{\delta(1-p)}e \qquad (7.60)$$

因此，为激励工人努力工作，企业必须向工人支付的工资不仅足以补偿工人自我雇佣时的机会收入以及努力工作带来的负效用w_0+e，还包含工资升水$\frac{1-\delta}{\delta(1-p)}e$。

若$p \to 1$（偷懒很难被发现），则工资升水必须非常高才可以激励工人努力工作。

若$p \to 0$，则当下式成立时工人努力工作才是最优的

$$w^* \ge w_0 + \left(1+\frac{1-\delta}{\delta}\right)e$$

式（7.60）只是在企业给定上述战略时，上述战略是工人的最优战略的条件。还须考虑给定工人战略时，冷酷战略是企业最优战略的条件。

给定工人的战略，企业在每一重复阶段的决策结 A、C、D 上选w^*（$\ge w_0 + \left(1+\frac{1-\delta}{\delta(1-p)}\right)e$），在 B 上及其后的决策结上都选$w=0$，这样规定了企业在其所有决策结上的行动选择，故给出了企业的战略，显然，因这样的w^*有无数个，这样的战略有无限多（但企业只会选$w^* = w_0 + \left(1+\frac{1-\delta}{\delta(1-p)}\right)e$）。其它可能的企业战略为：在 A、C、D 选$w < w_0 + \left(1+\frac{1-\delta}{\delta(1-p)}\right)e$（在 B 上不作规定），但因企业选冷酷战略$w^*$时的支付为$\frac{y-w^*}{1-\delta}$，选其它战略的支付为 0，因而冷酷战略为企业的最优战略的条件为$\frac{y-w^*}{1-\delta'} \ge 0$，$\delta'$是企业的贴现因子。

故有$y-w^* \ge 0$

$$y \ge w^*$$

因给定企业选冷酷战略时，已证工人选合作的条件为：

$$w^* \ge w_0 + e + \frac{1-\delta}{\delta(1-p)}e$$

故给定企业和工人的上述战略，它们构成纳什均衡的条件为：

$$y \ge w^* \ge w_0 + e + \frac{1-\delta}{\delta(1-p)}e$$

$$y - e \ge w_0 + \frac{1-\delta}{\delta(1-p)}e \qquad (5.63)$$

当且仅当式（7.62）满足时，均衡的w^*才存在。

因$\frac{1-\delta}{\delta} = \frac{1}{\delta} - 1$是$\delta$的减函数，故式（7.62）右边是$\delta$的减函数，所以，当贴现因子

δ充分大时（工人有足够耐心时），合作解才会出现。

子博弈精炼解证明

上面只证明了在一定条件下，上述战略组合是纳什均衡。下面证明它们还是子博弈精炼均衡。

有两类子博弈 { 始于高工资、高产出之后的子博弈(如图 7.15 中 A、C、D 开始)
其他子博弈(如图 7.15 中 B 之后的)

已证在第一类子博弈中，局中人的战略组合为纳什均衡，故只需看其他子博弈情形。

在其他子博弈中，给定企业选 $w=0$，工人选自我雇佣当然是最优的。反之，给定工人选自我雇佣，企业选 $w=0$ 为最优。

注意，在这一均衡中，自我雇佣是永远性的：即若工人曾有一次被捉住偷懒，则企业在其后将永远付给工资 $w=0$；若企业曾偏离 $w=w^*$，则工人将永不再努力工作，于是企业也不会再雇佣这个工人。

可以有许多理由对这种永远自我雇佣的合理性提出疑问。

譬如，在上述偏离后出现的永远性阶段博弈中，两个局中人是双败的，他们都希望重新开始无限次重复博弈并再度达成合作均衡（高工资、高产出），但这是不可能的。

因为，如果局中人事先知道惩罚将不会被执行，还有事后重新开始的可能，则惩罚就是不可置信的。

这就是为何在战略是相当“冷酷”的条件下，合作才会出现的原因。

在存在劳动市场情况下，若企业同时雇佣了许多工人，则它更倾向于不进行重新谈判，因为和一个工人的重新谈判会使处于高工资、高产出均衡的其他工人（或将要选择这一均衡的工人）十分失望而改变战略。

如果存在许多企业，问题成为企业 j 是否会雇佣以前企业 i 曾雇佣的工人。

合理的答案是“不会”。

因为它担心会使现有高工资、高产出的工人失望，正如一个企业的情况一样。

诸如此类的原因可以解释为什么日本大企业间白领男性的成年雇员缺乏流动性。

换一种情况，若被雇的工人总可找到比自我雇佣更喜欢的工作，则这时那些新工作的工资（减去努力带来的负效用）就起到了自我雇佣收入 w_0 的作用。

在一个被解雇工人根本不会受到任何损失的极端情况下，在无限重复博弈中无法提供对偷懒的有效惩罚，从而也不存在工人将努力工作的子博弈精炼纳什均衡。

事实上，Bulow 与 Rogoff（1989）曾给出一个例子：如果一个债务国能够在国际资本市场上通过预先收款的短期交易重复从债权国借入长期贷款，则在无限重复博弈中对债务国和债权国之间的违约行为就没有一个可行的惩罚方案。

存在一个问题是既然效率工资理论要求 $w^* > w_0$，而 w_0 又是工人在其它企业可获得的工资，若所有企业都实行效率工资，则又有 $w_0 > w^*$，岂不矛盾！这并不矛盾。当所有企业都实行效率工资时，给定一个工人，若他选偷懒，被发现开除后，到别的企业不能找到工作，因其它企业与所需工人间处于合作状态，所需工人够了，故事实上有 $w_0 = 0$。

注释：

[1] 这里假定局中人i的所有信息集即局中人i进行行动选择的信息集的集合是可以进行编号的，这种编号可以按局中人i的行动顺序的顺序来进行。

[2] 局中人i在其任一信息集上只能有唯一的一个行动空间，即对信息集上不同的决策结有相同的行动空间，否则，可以区别有不同行动空间的决策结，因而这些决策结不会位于同一个信息集内。

[3] 另二位获奖人分别是 Arkelof 和 Stiglitz。

[4] 这里假定了不同企业在观察到e时关于雇员类型的统计推断是相同的。

[5] 为了说明$L+R-I(1+r)<0$时债权融资是不可行的，我们来进行如下的分析。如果存在一种均衡，低盈利能力企业发出最优信号为D_L，当$D_L \geq I(1+r)$时，如果投资者接受，则企业支付为

$$L+R-D_L \leq L+R-I(1+r)<0$$

如果投资者不接受，则企业支付为$L+R>0$。显然，当D_L是最优信号时，投资者一定是拒绝的。这样，低盈利能力企业就不能通过债权方式为新项目融得资金。

[6] 当然，公众对于实际通胀率小于预期通胀率的意外情形似乎是感到欣慰的，此时可假设公众的单阶段支付函数为$\pi^e-\pi$。但是，一方面可以证明这种假定在增加了更多的数学复杂性情况下并未带来本质上不同的结果，另一方面，这种假定在经济学上可能是不妥的。因为，公众的效用最大化消费——工作计划或许是建立在其预期通胀率之上的，一旦实际通胀率偏离预期通胀率，则打乱了公众既定的消费—工作计划，其效用会降低，而不管这种偏离是$\pi>\pi^e$还是$\pi<\pi^e$。所以，用$-\left|\pi-\pi^e\right|$来测定其单阶段支付函数是有着充足理由的。但仅用绝对值作为支付函数可能造成数学分析上的不便，我们用预期通胀率与通胀率之间偏差的平方项来刻画偏差就方便多了。

[7] 在第二阶段，弱政府由式（7.37）制造较高的通胀率，这时它解除了伪装而暴露出真实的类型。

[8] 与前面的混同均衡一样，这里给出的只是充分条件而非必要条件，因为在非均衡路径上的信念有多种选择。

[9] 参见 Vickers(1986)对相似的两阶段博弈的详细分析，以及 Barro（1986）的多阶段声誉模型。

[10] 这一假设保证利润函数$M_1^{\theta}(P_1)$是单峰的，即只在一个点上达到极大。在此假定下，$P_m^H>P_m^L$，$M_1^H \subset M_1^L$是可以得到严格证明的。

[11]声誉模型只是假定理性人假定不是共同知识，并非放弃了理性人假定。即仍假定局中人都是理性的，但某些局中人不知道其他一些局中人到底是否理性，只知道其他一些局中人以某种概率判断为理性。这里，非理性不是指局中人不是效用最大化者，因

为这是博弈论的基本假定，而是指局中人可能的支付函数不是模型给出的，如囚徒困境中的某些囚徒的支付函数不是支付矩阵所给出的。即某些局中人关于其他局中人的支付函数是不完全信息的。

[12] 因非理性囚徒 1 的战略是给定的，不必通过最大化其支付导出，故不必提及非理性囚徒 1 的支付函数。

第八章　委托——代理理论：合约经济学

在博弈论不断取得进展的同时，经济学家们在一些应用领域分别独立地提出了后来被发现与博弈论有关的许多模型。这些来自不同的应用领域所产生的经济学问题，尽管几乎都是相互独立地被发现的，但后来被认为具有相同的理论背景，即几乎都是因为信息不对称导致局中人在不完全信息条件下进行博弈的结果。从经济分析的方法论角度看，这些模型实际上是博弈论在不完全信息条件下的应用。于是，人们就将这些原本是独立做出的模型集合成一个经济学的分支体系，这就是现在被人们所熟知的“信息经济学”体系。由于“信息经济学”是许多最初是独立做出的许多不同的模型的一个集合，所以，人们根据其不同的特征方面，也还给它取了其它的一些不同于“信息经济学”的名称，如委托—代理理论，合约理论（契约理论、合同理论）等等。

在 20 世纪的 80 年代，中国大陆曾组建了一个关注“信息经济学”的经济学家访问团赴美国斯坦福大学考察“信息经济学”的研究。当时，在斯坦福大学经济系任教的中国人钱颖一博士为该访问团约见了“信息经济学”的奠基人之一 K.Arrow 教授。但当中国大陆的“信息经济学家”们与 Arrow 教授座谈不到 10 分钟后，大家才发现双方在关于“什么是信息经济学”上难以沟通。原来，对中国大陆的“信息经济学家”们来说，此信息经济学非彼信息经济学也！在中国大陆当时的学术术语中，“信息经济学”是指研究诸如邮政、电讯等“信息产业”中的经济问题的学问，与国际上关于“信息经济学”的规范定义即研究不对称信息条件下经济行为的经济理论是不同的。由于这类可能出现的歧义，我们在本章以及下一章中将这种研究不对称信息环境中经济行为的理论称为含义更为具体的“委托—代理理论”。本章将介绍的内容限定在由 Mirrlees 及 Holmstrom 等人提供的静态博弈框架内，也称为合约理论。在下一章，将介绍 Fama 等人提出的动态博弈框架。

8.1 委托—代理关系：不对称信息

当你委托你的律师为你处理某件事务时，你与律师之间的委托—代理关系就发生了。在法律上，当某 A 授权于某 B 代表 A 从事某种活动时，就发生了“委托—代理”关系。其中 A 是“委托人”，B 是“代理人”。委托人与代理人的概念原初来自于法律学，但在经济学中，这种概念被进一步扩充到任何一种涉及非对称信息的交易活动，其中有私人信息的一方是代理人，而另一方就是委托人。在存在委托—代理关系的两个人之间，代理人的行为或代理人拥有的某些信息或知识不为委托人所能观察到或所知道，至少，委托人不能不花成本地观察到代理人的行为或掌握代理人拥有的某些信息。譬如，你的

律师或许没有尽心尽力地为你工作。这样，代理人就可能为了其自身的利益选择其行动而损害委托人的利益，称这种行为为代理人对委托人造成的“道德风险（Moral Hazard）。对于委托人来说，如果不是直接监督代理人的行为（这种监督是要花费成本的），就存在一个如何对代理人选择符合委托人利益的行动进行激励的问题。委托人可以与代理人签署合约，约定根据他们都能观察到的某些指标由委托人向代理人支付报酬或奖励的规则。委托人如何设计并挑选对其最有利的这类合约是委托—代理理论的核心问题，因而，委托—代理理论又被称为“合约（合同，契约）理论”。更为一般性的提法是，博弈的一个局中人如何设计对其最有利的博弈规划（如果这个局中人有能力支配其他局中人对其所设计的博弈规则的接受的话），也是一个更为广义的合约问题，称为“机制设计”问题。所以，合约理论、委托—代理理论、信息经济学、机制设计等等，这些不同的名称通常说的是一回事。

在你委托律师为你工作的例子里，不能无成本地观察到的行动或掌握的信息发生在你与律师发生“委托—代理关系”之后，称为“事后”的信息不对称。与之相对应，发生在“事前”的信息不对称包括诸如招聘、投保、商品销售和银行贷款的过程中。当一家公司招聘员工或经理时，对于应聘者的真实能力的信息，公司是不清楚的。投保人向保险公司投保时，如果是人身保险或寿险，保险公司也拿不准投保人身体健康状况的真实情况。消费者在购买商品时，不能掌握有关商品质量的全部信息。商业银行在为一个项目贷款时，也一般不能对项目的风险有完全准确的估计。

类似地，发生在“事前”的信息不对称还包括信息本身的不对称，称为“知识”的不对称。这是指一个局中人所掌握的某些知识不被另外的局中人在事前所知。对于这样的一些有关信息不对称的场合，我们可以举出下面的一些例子。

为了将众多的信息不对称例子加以归类，我们将分别发生在事前和事后，以及是关于行动还是关于信息或知识的信息不对称分为以下几种：

8.1.1 事前的隐藏信息博弈

这类博弈包括逆向选择（Adverse Selection）、信号传递（Signaling）、信息甄别（Screening）。

逆向选择的著名例子是二手车市场；信号传递就是信号博弈；其例子有第七章介绍的 Spence 模型；信息甄别是一种解决事前信息不对称的机制设计，它是通过分离均衡而达到将不同类型局中人加以识别开来的目的。在这类不对称信息博弈中，还包括保险市场、金融市场、垄断者价格岐视、公司内部持股比例、公司资本结构等模型。在保险公司与投保人之间签订保险合约时，保险公司不是很清楚投保人的健康状况。商业银行

在贷款给企业时，对企业或项目的还款能力也不是很清楚的。垄断者在销售其产品时，不是很清楚顾客的需求强度，因而设计一些岐视性价格来揭示出顾客的需求强度类型，此时博弈表现为信息甄别。公司内部持股比例愈高，说明公司愈好，因为内部人比外部投资者更清楚公司的实力，这也是一种信号传递博弈。正如第七章中的 Ross 模型所表达的那样，公司资本结构也会向外部投资者发送有关公司实力的信号，这是事前信息不对称的一种解决方案。

8.1.2 事后的隐藏信息博弈

这是一种道德风险模型所表达的情形。在这类博弈中，有股东与经理之间、债权人与债务人之间、经理与销售人员之间、雇主与雇员之间、原告或被告与代理律师之间的委托—代理关系。经理作为股东的代理人，可能会做出利己但损害股东利益的道德风险行为。债务人可能将债权人借给他的钱用于高风险项目，从而损害债权人利益。销售人员可能未尽心尽力推销企业产品，但又将不良的销售业绩归咎于市场需求不足等等。

8.1.3 事后的隐藏行动博弈

这也是一类道德风险模型描述的情形。当投保人在取得保险合约之后，不保重身体（不良生活习惯如饮酒、吸烟等），或不注意防盗、不注意汽车保养、佃农不努力劳作、经理不努力经营、雇员不努力工作、债务人不控制项目风险、房东不加强房屋修缮、房客不注意房屋维护、议员不真正代表选民利益、政府官员不廉洁奉公、律师不努力办案时，事后隐藏行动的道德风险就出现了。对一个社会来说，犯罪分子的犯罪行为也是这样的一种道德风险。

总的看来，非对称信息可按时间在“事前”和“事后”发生的可能性分为事前非对称和事后非对称，也可按内容上的非对称分为“行动上的非对称”和“知识上的非对称”，分别称为“隐藏行动”和“隐藏信息”的博弈。

对于道德风险，是因为委托人不能完全观察到代理人的行为，而代理人活动的结果尽管能被观察到，但这种结果不完全是代理人行动选择的结果，而是代理人行动与其它的随机性因素共同作用的结果。并且，委托人不能将代理人行动与随机因素的作用完全区分开来。如土地上的农作物产量是佃农努力工作程度与随机性的气候条件共同作用的结果，地主并不能将佃农的贡献与随机性的气候条件的贡献分开来。

逆向选择、信号传递及信息甄别博弈实际上都是事前的信息不对称环境下的博弈，后两者是解决逆向选择问题的机制设计。信号传递和信息甄别机制在解决逆向选择问题时是相似的。

需要指出的是，同一个委托—代理关系可以存在多种的信息不对称属性，如雇主知道雇员的能力但不知其努力水平时，是一个隐藏行动的道德风险问题；但若雇主和雇员本人在签约时都不知道雇员的能力，但雇员本人在签约后发现了自己的能力（雇主仍不知），则问题就是一个隐藏信息的道德风险问题。若雇员开始就知道自己的能力而雇主不知，则是逆向选择问题，若雇员开始就知道自己的能力而雇主不知且若雇员在签约前就获得学历证书，则问题就是信号传递问题。相反，若雇员是在签约后根据工资合同的要求去接受教育，则问题就是信息甄别问题。

某些经济学家认为隐藏信息的道德风险博弈和信息甄别及逆向选择博弈在本质上是一回事。如 Myerson（1991，P263）提出将所有“由局中人选择错误行动引起的问题”称为“道德风险”，而所有“由局中人错误报告信息引起的问题”称为“逆向选择”。一些经济学家认为信号传递博弈与信息甄别博弈没有多大区别，因而也可用“信号传递”来指两种情况。

严格说来，可用“委托—代理”来概括上述五种类型的博弈，但在习惯上，“委托—代理”通常指“隐藏行动的道德风险博弈”。在本书中，我们仍然遵守这一习惯，除非有特别的说明。

这样，我们可以将所有的信息经济学模型分为两类，即委托—代理模型和逆向选择模型（包括了信号传递和信息甄别模型）。

信息经济学与博弈论之间的关系是，前者是后者在信息不对称环境下的应用，但从特点上看，博弈论更注重于方法论，而信息经济学注重于问题的解析。博弈论研究的是给定信息结构下的均衡是什么，而信息经济学研究的是给定信息结构下，什么是最优的合约安排。

信息经济学主要研究非对称信息环境中的最优合约，故又称为合约（契约，合同）理论或机制设计理论。

博弈论从某种意义上看是“实证的”，而信息经济学是“规范的”。

委托—代理理论在经 Fama 等人的发展后引入了代理人市场竞争的因素，从而在理论的结构上有新的发展。本章将介绍在 Fama 的发展之前的理论框架，我们称之为合约理论，而将 Fama 等人的体系放在后面的第九章中去介绍。

8.2 委托—代理：分析的框架

8.2.1 基本思想

在本节中，我们来设计一个可用于分析委托—代理关系的基本框架。就委托—代理

的一般意义来说，所谓委托—代理是指委托人通过给予代理人一定的奖赏去诱使代理人按照委托人的利益要求完成一定的行为。在这个非常一般的理解中，其实包含了十分丰富的内容。首先，委托人给予代理人的奖赏无论在形式还是在数额上都是有多种可能的。奖励可分为物质和精神上的，甚至还可能包括诸如权力在内的奖赏内容。代理人需要完成的行为既包括"事后"的，也包括"事前"的。"行为"本身也包括"行动"和"知识报告"等多种内容，正如前面一节中所分析的例子所指出的那样。在本章，我们将奖赏限制在物质的内容上，将"行为"限制在"行动"上，并且是"事后"发生的。当然，按照本章的一贯假定，委托人不能无成本地观察到代理人的行为，即代理人的行为是"私人信息"。

在代理人行为是委托人不能无成本地观察到的情形，委托人就面临代理人说谎的风险。一个佃农可以将因其偷懒而造成的产量下降归咎于不利的气候；国有企业的经理也可能将亏损归咎于过去发生的大量负债或职工们在大锅饭分配制度中的懈怠（尽管这类因素确已构成国有企业不振的一部分原因）。此时，委托人存在两种可选择的方式方式去处理这种风险。一是直接去观察代理人的行为，譬如，老板雇用监工去监视工人的劳动。但是，直接观察是要花费额外成本的，如老板要为监工开工资。另外，老板雇用监工实际上又引入了新的一种委托—代理关系，即老板作为委托人请监工代理其监督工人的工作。这样，又存在监工的工作是否努力的问题，是否还需要再雇用监工的监工去监视监工呢？显然，除非老板自己直接去监视工人，委托—代理关系及其带来的因代理人行为不可无成本观察的问题都会对委托人带来一种额外的成本。但是，即使由老板自己亲力而为地去监视工人，老板也会花费另一种成本，即老板自己的时间机会成本，因为倘若老板将用于监视工人的时间作其它用途（制定战略计划、营销或休闲），他会获得其它的效用。我们可以假定，随着监视的增加（监视所花费的时间或成本的增加），监视的边际成本是递增的，而监视的边际收益下降。这样，总存在代理人的一些"剩余"行为，倘若老板通过监视去观察这些行为，其边际收益会小于边际成本。此时，再通过边际上的监视的增加（监视的时间的增加）去观察代理人行为就是得不偿失的。对于这样的场合，即监视的成本大于收益的情形。委托人再通过观察代理人行为去控制因代理人偷懒的风险就是不经济的了。此时，存在另外的一种可选择的方式就是：委托人与代理人之间签订合约。

当代理人的行为不能通过"经济"（即观察成本小于观察收益的情形）的方法被委托人观察到时[1]，委托人与代理人之间就会就代理人的"真实行为"产生分歧。代理人会利用委托人不清楚代理人行为的真实状况这一点向委托人提交有利于自己，而不利于委托人的报告。譬如，佃农会向地主夸大其工作努力的程度。当然，委托人也会因其不

能观察到代理人的行为而难以相信代理人的一面之词，并且，他也知道代理人有说谎的动机。解决这一问题的一种办法是：合约将委托人支付给代理人的奖赏与某个委托人与代理人都同时承认可以共同观察到的指标相联系起来。并且，这种指标的可观察性是共同知识，即委托人和代理人都能观察到这种指标，委托人和代理人都知道对方能观察这种指标，委托人和代理人都知道对方知道自己能观察到这种指标……。所谓能观察到的指标，是指观察成本足够小，我们这里将观察成本假定为零。当然，这种指标可以不止一个，如企业把考核员工的工作绩效用某个包含有多个单一指标的“指标体系”来“测算”。我们在本章的后面部分，将分析什么样的的指标应进入“指标体系”，或一个“指标体系”究竟应包括多少指标的原理。

现在的问题是，尽管委托人和代理人就某一指标的“可观测”性上可以达成共识，但可能委托人和代理人就指标的“预测值”发生分岐。如国有企业的政府主管部门与国有企业经理就企业的产出产生不一致的说法。这种问题在原则上是完全可能出现的，正如 Hart （Oliver Hart, 1995）在其“不完全合约理论”（Uncomplete Contract theory）中所指出的那样。我们可以假设存在一个双方都接受的第三者，根据第三者提交的观测来确认指标的预测值。当然，正如不完全合约理论还进一步指出的那样，由谁来充当第三者自然是一个没有解决的问题。我们在这里为避开“不完全合约”理论所揭示出来的复杂性，假定这个困难是不存在的，如请政府指定的审计部门提交有关国有企业的利润指标。[2]

从博弈论看，这种根据某个双方都共同预测到的“指标”来决定委托人和代理人支付的合约，实际上是一种“相关博弈”。指标就是“信号”，而委托人支付给代理人的奖赏与代理人因工作辛苦形成的成本共同决定代理人的支付。

委托人的问题是，选择什么样的“指标（体系）”，以及什么样的合约，使委托人感到最为满意（效用最大化）？委托—代理理论的中心问题决定要为此寻找出一个答案，所以，委托—代理理论又可称为“最优合约理论”或“最优合同理论”。

8.2.2 委托—代理理论的起源

从学术思想的历史回溯，合约经济学的最初想法来自张五常在其博士论文《佃农理论》中所作出的分析。

1965 年，在美国加州长堤大学教书的张五常为了完成他的博士论文，在图书馆查阅资料时，偶然发现了一些有关台湾农业的资料，台湾土地改革的一些情况随即引起了他的浓厚兴趣。

地主与佃农之间的合约类型有三种，一是地主获得固定的租，而剩余给予佃农；二

是地主获得剩余，佃农获得固定的收入；三是地主与佃农之间按固定的比例获得土地上的产出。在第一种情形，佃农承担了全部风险，而地主获得无风险的收入；第二种情形相反，地主承担全部风险而佃农获得固定的无风险收入；第三种情形地主与佃农之间分摊了风险。由于农业生产的风险很大，因为产出受不可预知的气候变动的影响很大。在第一种情形，佃农处于很不利的地位，而中国在 1949 年之前的漫长历史中，土地与佃农之间的合约大概都是以这种类型为代表的。所以，当国民党政权在大陆崩溃后，他们在台湾就开始反省自己失败的原因，结论是他们在大陆失去了农民的支持，而不利于农民的土地租约显然是其原因之一。于是，国民党政权在台湾开始了 20 世纪 50 年代的土地改革，方法是改变传统的让农民承担全部风险的土地租约，代之以让地主与佃农之间分摊风险的第三类租约，即在地主与农民之间实行土地佃营的“分成制”。

分成租佃问题，古典经济学家（如亚当·斯密）及新古典经济学家（如马歇尔）都有所关注，但是在以往的佃农理论看来，佃农在分成租佃制下，劳动者的劳动收益都会小于其边际产出，因此，这种制度安排是低效率的或无效率的。但是，在欧洲，分成制长期都存在。人们自然要问，一种低效率或无效率的制度安排为什么能够长期存在并长盛不衰呢？人们为什么不创立新的制度安排来取代它？张五常研究了台湾 20 世纪 50 年代期间的土地改革，并从中引伸出了一般性的理论。

台湾 1949 年开始了土地改革。在改革中，当局把地主与农民的分成，规定为地主的地租不得超过农产品的 37.5%。但让经济学家难以置信的是，就在当局的这种硬性约束下，台湾农产量非但没有如斯密和马歇尔理性所预言的那样下降，反而出现了急剧的上升。这是不是台湾当局搞的数字游戏？张五常带着这个疑问进行了多方的调查，但调查的结果却显示，台湾土地改革后农业产量上升的数字是无可置疑的。

为什么台湾当局对地租进行管制，生产没有下降反而却上升了呢？

张五常这时便尝试着用产权理论来解释这种情况：首先，土地是地主的私有财产，劳动力是农民的私有财产，所以要从私有财产的局限性入手对此进行解释；其次，佃农分成制是一种合约，它与任何合约一样，其中的条件是由双方议定的；再次，农民之间不可避免地要展开竞争，地主之间也有竞争，因此佃农合约中的条件（包括分成的百分比），是在私有财产竞争的这两种局限下决定的。所以在分成制下，农民与地主的投资，佃农租种土地面积的大小与耕种劳力的多少，以及地主与农民分成的百分比，都是由上述三个理论基础决定的。这些被决定的项目，就是佃农合约的条件。

从上述理论出发，张五常很容易地推断地主在土地上的分成收入与固定租金、雇用农民、自耕自种等不同形式的收入大致相同。因为，竞争的局限条件是大体一样的。在这个思路的基础上，张五常又进一步研究了在资源的运用与收入的分配大致相同的情况

下，为什么会有不同的合约安排的问题，从而导致了在随后大行其道的合约经济学的研究。在上述理论做得满意以后，张五常又加入台湾土地改革中的约束地主分成的百分比。于是，他很快就把台湾土地改革中，当局对地租施加限制，但农产量不降反增这个问题从理论上解释清楚了：这是因为这个约束是在市场决定的地主分成之下，农民的收入会高于他们另谋高就的收入，所以在竞争下他们必须努力工作，使地主在较低的分成率中因为生产的增加而有点补偿。这样一来，农产量自然就上升了，这就是张五常独特的佃农理论。

张五常的研究给后来的委托—代理理论带来的启发包括：如果代理人不承担任何风险，如佃农取得固定的收入，则由于代理人的行为不能“经济”地被委托人观测到，代理人就必定会偷懒并向委托人谎报其努力程度。因此，代理人必定要承担一定的风险，他才会有努力工作的动机。这是因为，当代理人承担一定风险时，其收入是随产出的增加而增加的，倘若完全不努力，他不会有收入的。但是，我们在考虑风险收入的场合，需要考察在个人的效用函数中引入风险变量。如果代理人惧怕风险（后面将要定义的风险规避型）。则只有当代理人在收入上获得一定的额外补偿时，代理人才愿接受风险收入合约。这种额外的补偿来自委托人的收入，所以，倘若让代理人承担全部风险，则委托人需要支付过多的收入对代理人进行补偿，这也许对委托人是不经济的。因而，可以事先预见到的是，一般情况下的合约应该是委托人与代理人都同时承担一定的风险，而这正是“分成制”合约。在台湾的土地改革中，地主分成的比例存在上限限制，且上限还比较低，这就给予了佃农以较大的风险补偿，以鼓励农民选择种地，这在后面的概念框架中被称为“参与约束”。[3]当地主和佃农都各自存在竞争时，均衡状态下的土地报酬率是可以预见的，因而毋须地主去观测佃农的努力程度，地主就应判断出佃农是否足够努力。这在委托—代理理论中正是“标尺竞争”的含义。张五常在其博士论文中写道：

事实上，在私有产权条件下，给定了土地可自由转让（可市场化）的权利后，一个土地所有者可以不必亲自了解农业耕作的细节，对资源所有权的竞争就会诱导出一个有效的合约。如果佃农种植的作物价值不大，如果租金率太低，而每一个佃农所承担的土地规模太大，或所要求的佃农投入太少，则作为给予土地所有者的土地报酬的年租金率将低于利息率。在这种情况下，土地所有者可能会进行适当的调整，把土地租给其他的佃农，从而选择一个不同的合约安排，他可直接出卖土地所有权。另一方面，如果合约的安排是佃农的分成收入低于他的可供选择的收入，则其他的土地所有者会出高价获得他的劳动。

由于土地所有者之间存在着竞争，佃农之间的竞争会确保他们完成合约所规定的投入量。事实上，土地所有者只需要通过检查产出就可以知道佃农是否遵守合约条款，从

而决定分成合约是否能继续下去（张五常，1968）。

8.2.3 数学表述

从经济思想上看，可以说张五常在其博士论文《佃农理论》中已经将委托—代理的理论要旨表达出来。但是，作为建立在严格的公理体系中的委托—代理理论，张五常并未建立起来，这要等到后来的博弈论经济学家们的介入，才最终将严谨的理论构架描绘出来。委托—代理理论的数学框架，最早应归于 Hurwitz 教授的早期工作。但现在我们在本章中将主要给出的框架，主要是牛津大学和芝加哥大学的学们者的工作。

假设代理人可以选择的行动的集合为 A，$a \in A$ 是代理人可能选择的一个行动。a 可以是一个向量，如当 $a=(a_1,a_2)$ 时，a_1 是“数量”、a_2 是“质量”。在本章中，我们将 a 限制在是一维变量的情形。

设 θ 是外生的随机变量，它是不受代理人和委托人控制的“自然状态”，$\theta \in H$，H 是 θ 的可能取值范围，并设 θ 在 H 上的分布函数和密度函数分别为 $G(\theta)$ 和 $g(\theta)$。

当代理人选择某个具体的行动 a 之后，外生变量 θ 实现，θ 与 a 就共同解决了一个可预测结果，记为 $x=x(a,\theta)$。我们进一步还假设 a 与 θ 会共同决定一个所有权归属于委托人的货币收入（产出）$\pi(a,\theta)$。$x(a,\theta)$ 也可以是一个向量，此时它就是我们在前面所说的“指标体系”。$x(a,\theta)$ 也可以将 π 甚至 a 和 θ 都作为它的分量。当 a 或 θ 是 x 的分量时，a 就是可观测的了。

委托人和代理人在收入不确定情况下的效用函数并不简单地等同于他们在确定性情形下的效用函数。设委托人的确定性收入效用函数为 $v=v(y)$，其中 v 是委托人在收入为 y 下的效用水平，又设 $u=u(z)$ 是代理人的确定性收入效用函数，其中 u 是代理人在收入为 z 下的效用水平。假设这些效用函数满足通常的性质假定，即 $v'=\dfrac{dv(y)}{dy}>0$，$v''=\dfrac{dv}{dy}\le 0$；$u'=\dfrac{dv(z)}{dz}>0$，$u''=\dfrac{du'}{dz}\le 0$。

当确定性效用函数给定时，如何构造其在不确定性情形的效用函数是经济学在目前也未解决的一个难题。这是因为，按照经济学的现代理论，效用函数是序数的，并在任何严格单调递增函数的复合下也是同一编好序的效用函数。但我们至今也未能构造出一种不确定性效用函数，它既满足这一复合性质，又满足当不确定性情形退化到确定性情形时，不确定性效用函数也正好回到给定的确定性效用函数。尽管如此，Von Neumann 与 Morgenstern（1944）证明，期望效用函数能满足当复合的函数仅限于正线性变换函数的情形保持同一偏好序效用函数的性质。也就是说，当期望效用函数 $\int v(y)f(y)dy$（其中 $f(y)$ 是随机收入 y 的分布密度函数）被一个正线性变换函数 $z=cv+d$（其中 c，d 为

常数，$c>0$）复合成$z=c\int v(y)dy+d$，则z也是委托人的一个效用函数。同时，当不确定性退化到确定性情形$y=y_0$时，即$f(y)=\delta(y-y_0)$，其中$\delta(y-y_0)$是 Dirac 函数，满足

$$\delta(y-y_0)=\begin{cases}+\infty & y=y_0\\ 0 & 其它\end{cases}$$

且$\int\delta(y-y_0)dy=1$，此时有

$$\int v(y)f(y)dy=v(y_0)$$

这是一个“次优”的结果，称期望效用函数为 V Neumann-Morgenstern 效用函数或 V-N-M 期望效用函数。该函数目前在经济学中处理不确定分析时几乎是唯一被使用的不确定性效用函数。

代理人选择任何行动a几乎都会给他带来一定程度的“辛苦”或“痛苦”，假定其可被用一种效用测度的“成本函数”$c(a)$来刻画，且假设有$c'=\dfrac{dc(a)}{da}>0$，$c''=\dfrac{dc'}{da}>0$。

显然，一般可假定$\dfrac{\partial\pi(a,\theta)}{\partial a}>0$。这与$c'>0$构成一对矛盾。$\dfrac{\partial\pi}{\partial a}>0$意味着委托人希望代理人多加努力，而$c'>0$则意味着代理人希望少努力。所以，除非委托人能对代理人提供足够多的激励或奖赏，否则，代理人不会如委托人希望的那样努力工作。

假设分布函数$G(\theta)$（密度函数$g(\theta)$），生产技术$x(a,\theta)$，产出函数$\pi(a,\theta)$、效用函数$v(y),u(z)$，成本函数（也称“负效用函数”）$c(a)$都是“共同知识”。[4]

显然，$x(a,\theta)$是共同知识时，当委托人能观测到θ时，也就可以知道a，反之亦然。

所以，我们一般总假定a和θ同时都是不可观测的。下面介绍参予约束的概念，参与约束（个人理性约束）：

设委托—代理关系为：委托人将产出$\pi(a,\theta)$中的一个部分s作为奖赏支付给代理人。并且，合约规定，s是按照可观测变量（指标）$x(a,\theta)$来决定的，即有

$$s=s[x(a,\theta)]$$

则委托人的收入为$y=\pi(a,\theta)-s[x(a,\theta)]$，于是，委托人的 V-N-M 期望效用函数为

$$(P)\qquad \int v\{\pi(a,\theta)-s[x(a,\theta)]\}g(\theta)d\theta$$

委托人的问题是选择$s(x)$以及a（a是由代理人选择的，但委托人的问题是他希望什么样的a）使这一期望效用函数最大化[5]。

代理人不参加与委托人的这一委托—代理博弈即不签订合约时，他也会有一个“保留支付”或“保留效用”，记为$\bar{u}$。它是代理人不接受合约时的最大期望效用，即代理人接受合约的机会成本。

于是，代理人的净期望效用函数为

$$(IR) \qquad \int u\{s[x(a,\theta)]\}g(\theta)d\theta - c(a) \geq \bar{u} \qquad （8.1）$$

式（8.1）称为“参与约束”或“个人理性约束”（individual rationality constraint，称写为 IR），它是代理人接受合约的必要条件。

另一个概念是激励相容约束，下面给出定义。

激励相容约束：

尽管委托人不能“经济地”观测到代理人的行为，但有一个原理制约着代理人的行为，这就是“激励相容约束”（incentive compatibility constraint，简写为 IC）。这个约束决定了代理人的行动选择a应满足的条件：

$$(IC) \quad \begin{aligned} &\int u\{s[x(a,\theta)]\}g(\theta)d\theta - c(a) \\ &\geq \int u\{s[x(a',\theta)]\}g(\theta)d\theta - c(a'), \forall a' \in A \end{aligned} \qquad （8.2）$$

一个委托—代理博弈中，委托人应清楚代理人的行动选择a必须同时满足(IR)和(IC)这两个约束。

这样，委托人的问题就是：在(IR)和(IC)限定的范围内选择a（通过奖惩诱使代理人选择a）和$s(x)$，最大化期望效用函数(P)，即：

$$\begin{aligned} &\max_{a,s(x)} \int v\{\pi(a,\theta) - s[x(a,\theta)]\}g(\theta)d\theta \\ &s.t \quad (IR) \quad \int u\{s[x(a,\theta)]\}g(\theta)d\theta - c(a) \geq \bar{u} \\ &\qquad (IC) \quad \int u\{s[x(a,\theta)]\}g(\theta)d\theta - c(a) \\ &\qquad \geq \int u\{s[x(a',\theta)]\}g(\theta)d\theta - c(a'), \forall a' \in A \end{aligned} \qquad （8.3）$$

这就是由 wilson（1969），Spence 和 Zeckhauser（1971）及 Ross（1973）等人提出的委托—代理博弈的分析框架，称为“状态空间模型化方法”（State-Space formulation）。

这种方法的优点是每一种技术关系都很直观地表达出来，但困难的是由该方法难以导出有信息量的解（若$s(x)$不限制在一个有限的区域，这个模型还可能没有解）。

在本书中将主要使用的方法是由 Mirrlees（1974、1976）和 Holmstrom（1979）提出的所谓“分布函数的参数化方法”（Parameterized distribution formulation）。

这种方法的基本思路是：

因为$x = x(a,\theta)$，$\pi = \pi(a,\theta)$，所以，对于每一个固定的a，θ与x或者θ与π是相对应的。因为θ是随机变量，故此时x和π都是随机变量。

我们将θ的分布函数转换为x和π的联合分布函数，用$F(x,\pi,a)$和$f(x,\pi,a)$分别代表从分布函数$G(\theta)$导出的联合分布函数和密度函数。

此时，委托人的问题就可表示为：

$$\max_{a,s(x)} \int v\{\pi - s(x)\}f(x,\pi,a)dx \qquad （8.4）$$

$$s.t \quad (IR) \ \int u[s(x)]f(x,\pi,a)dx - c(a) \geq \bar{u}$$

$$(IC) \ \int u[s(x)]f(x,\pi,a)dx - c(a) \geq \int u[s(x)]f(x,\pi,a)dx - c(a'), \forall a' \in A$$

除了上述两种方法之外，还有一种更加抽象的分析框架是所谓的一般化分布方法（general distribution formulation）。这种方法基于在分布函数的参数化方法表述下，代理人选择不同的行动 a 等价于他选择了不同的分布函数 $F(x,\pi,a)$（或不同的密度函数 $f(x,\pi,a)$）。

由此，我们就可将分布函数本身当作代理人的选择变量，从而将 a 消掉了（用 F 或 f 对应于 a）。

设 p 是 x 和 π 的一个密度函数，P 为所有 p 的集合 $(p \in P)$。因 $c(a) = c[a(p)] = c(p)$（由 p 与 a 的上述对应），故 $c(p)$ 为 p（对应某个 a）的成本（负效用）函数。则委托人问题又可表述为：

$$\max_{p \in P, s(x)} \int v\{\pi - s(x)\} p(x,\pi)dx$$

$$s.t \quad (IR) \ \int u[s(x)]p(x,\pi)dx - c(p) \geq \bar{u} \qquad （8.5）$$

$$(IC) \ \int u[s(x)]p(x,\pi)dx - c(p) \geq \int u[s(x)]\tilde{p}(x,\pi)dx - c(\tilde{p}), \forall \tilde{p} \in P$$

在这种表述中，关于 a 和成本 $c(p)$ 的经济学解释消失了，但得到一个非常简练的一般化模型，这个一般化模型甚至包括隐藏信息模型。

在上述三种表述方法中，参数化方法是标准的方法，本章将主要采用这种方法。

今后将假定产出 π 是唯一的可观测指标（即 $x = \pi$）。委托人对代理人的奖惩只能根据观测到的产出 π 作出。这时，委托人的问题就是：

$$\max_{a, s(x)} \int v\{\pi - s(x)\} f(\pi,a)d\pi$$

$$s.t \quad (IR) \ \int u[s(\pi)]f(\pi,a)d\pi - c(a) \geq \bar{u} \qquad （8.6）$$

$$(IC) \ \int u[s(\pi)]f(\pi,a)d\pi - c(a) \geq \int u[s(x)]f(\pi,a')d\pi - c(a'), \forall a' \in A$$

8.3 最优合约：对称信息的情形

8.3.1 直观推理

委托—代理问题的核心在于由于信息不对称带来的激励问题。也就是说，由于委托人不能无成本地或“经济地”观测到代理人的行为或把握代理人知道的知识状态，委托人就需要设计有效的合约来激励代理人最大程度地按照符合委托人利益选择其行为。但是，为了对比信息对称情形与信息不对称情形下的博弈结果，从而刻画由于信息不对称带来的交易成本，我们在本小节中暂时不考虑信息的不对称，假定委托人可以无成本地

观测到代理人的行动。

我们先来看看在这种信息对称情形下的博弈结果是什么，并将其作为后面研究信息不对称情形下的博弈的一个参照。

在委托人可以无成本地观测到代理人的行动的情况下，委托人的最优合约是什么？代理人的行动选择是什么？在给定合约下，委托人希望代理人选择什么样的行动？委托人如何设计博弈机制使代理人选择委托人希望他选择的行动？这些问题，都是以下分析将要回答的。

在信息对称的情形，分析委托—代理问题不需要激励相容约束，于是，委托人的问题是：

$$\max_{a,s(\pi)} \int v\{\pi - s(\pi)\} f(\pi, a) d\pi \qquad (8.7)$$

$$s.t \quad (IR) \quad \int u[s(\pi)] f(\pi, a) d\pi - c(a) \ge \bar{u}$$

显然，给定合约 $s(\pi)$，委托人不会希望代理人无限大地努力。这是因为，尽管代理人愈努力，即 a 愈大时，产出 π 的预期水平将愈高。但是，随着 a 的增大，代理人的负效用即成本 $c(a)$ 也会增加。由于参与约束，委托人要让代理人参与博弈，就需要给予代理人愈多的奖赏。由于边际产出的递减性质和边际成本函数的递增性质，必定存在某个有限的努力程度 a^*，使委托人的期望效用达到最大化。即委托人在给定合约下最希望代理人选择的努力程度是有限的。

那么，这个有限的努力程度 a 是如何能被加以数学刻画的呢？

显然，在委托人的期望效用达到最大的情形，参与约束 (IR) 必定会变成等式，因为他不必在已能诱使代理人参与博弈的情况下还付给代理人额外的产出。

当代理人的行动选择 a 已位于委托人最希望他选择的水平 a^* 时，在边际上，代理人的努力程度再无穷小地增加 1 个单位时，其成本将增加 $\frac{\partial c(a)}{\partial a}$。根据参与约束，委托人需要额外支付给代理人的最小（能保证代理人还参与博弈）收入为 $\frac{\frac{\partial c(a)}{\partial a}}{u'}$，其中 $u' = \frac{\partial u(s(\pi))}{\partial s}$。这是因为，在边际上，代理人收入的无穷小 1 单位增加将带来的效用增加为 u'，而边际上增加的效用 $\frac{\partial c(a)}{\partial a}$ 是 u' 的 $\frac{\partial c(a)}{\partial a}\Big/u'$ 倍，故代理人收入的无穷小增加为 $\frac{\partial c(a)}{\partial a}\Big/u'$ 个单位。

这种给予代理人的收入补偿来自委托人收入的减少部分，它给委托人在边际上带来的效用减少为 $v'\frac{\partial c(a)}{\partial a}\Big/u' = \frac{v'}{u'}\frac{\partial c(a)}{\partial a}$，其中 $v' = \frac{\partial v(\pi - s(\pi))}{\partial y}$，$y = \pi - s(\pi)$。这是在给定

的不确定因素下的分析，由于θ是随机变化的，代理人的期望效用减少为$\int \frac{v'}{u}\frac{\partial(a)}{\partial a}g(\theta)d\theta$。

当代理人努力程度a在a^*上有一个无穷小的 1 单位增加时，给定θ下产出增加为$\frac{\partial\pi(a,\theta)}{\partial a}$，委托人效用增加为$\frac{v'\partial\pi(a,\theta)}{\partial a}$，期望效用增加$=\int v'\frac{\partial\pi(a,\theta)}{\partial a}g(\theta)d\theta$。我们以下用“期望算子”$E$来表示对一个随机变量计算期望值，即

$$Ev'\frac{\partial\pi(a,\theta)}{\partial a}=\int v'\frac{\partial\pi(a,\theta)}{\partial a}g(\theta)d\theta$$

$$E\frac{v'}{u'}\frac{\partial c(a)}{\partial a}=\int\frac{v'}{u'}\frac{\partial c(a)}{\partial a}g(\theta)d\theta$$

在最优努力程度a^*处，根据经济学原理有

$$E\frac{\partial\pi(a,\theta)}{\partial a}=E\frac{v'}{u'}\frac{\partial c(a)}{\partial a} \tag{8.8}$$

式（8.8）就是通常的“边际成本等于边际收益”法则的一种变体，即委托人在“边际效用的减少等于边际效用的增加”处达到期望效用的最大化。

8.3.2 严格的数学推导

下面，我们通过严格的数学推导来获得式（8.8）。

构造问题（8.7）的拉格朗日函数[6]

$$\begin{aligned}L[a,s(\pi)]=&\int v\{\pi(a,\theta)-s[\pi(a,\theta)\}g(\theta)d\theta\\&+\lambda\left\{\int u[s(\pi(a,\theta))]g(\theta)d\theta-c(a)-\bar{u}\right\}\end{aligned}$$

其中，λ是拉格朗日乘数。

最优化的一阶条件为：

$$\frac{\partial L}{\partial s}=\int[-v'+\lambda u']g(\theta)d\theta=0 \tag{8.9}$$

$$\frac{\partial L}{\partial a}=\int v'\left[\frac{\partial\pi}{\partial a}-s'\frac{\partial\pi}{\partial a}\right]g(\theta)d\theta+\lambda\left[\int u's'\frac{\partial\pi}{\partial a}g(\theta)d\theta-\frac{\partial c(a)}{\partial a}\right]=0$$

如果委托人不知道$g(\theta)$，但这与前面假定的$g(\theta)$是委托人和代理人都知道的“共同知识”的假定相矛盾。倘若我们放弃$g(\theta)$是委托人知道的一个函数的假定（这是可能的，这意味着随机因素θ对于委托人来说不仅是随机的，而且还是“不确定的”。不确定指随机变量的概率分布都是未知的情形），则由 Wellstrass 定理知由式（8.9）可得到：

$$-v'+\lambda u'=0 \tag{8.10}$$

或者，我们仍然坚持$g(\theta)$是“共同知识”，但这意味着（8.10）只是一个充分条件。[7]

式（8.9）的第二个式子是a^*满足的一阶条件，将式（8.10）代入（即$\lambda=\frac{v'}{u'}$），则

式（8.9）的第二式为：

$$\int v'\left[\left(\frac{\partial \pi}{\partial a}-\frac{\partial s}{\partial \pi}\frac{\partial \pi}{\partial a}\right)g(\theta)+\frac{\partial s}{\partial \pi}\frac{\partial \pi}{\partial a}g(\theta)\right]d\theta-\lambda\frac{\partial c(a)}{\partial a}=0$$

即

$$\int v'\frac{\partial \pi}{\partial a}g(\theta)d\theta-\lambda\frac{\partial c(a)}{\partial a}=0$$

或者写为期望算子形式：

$$E\left[v'\frac{\partial \pi}{\partial a}\right]=E\left[\frac{v'}{u'}\frac{\partial c(a)}{\partial a}\right]$$

这正是式（8.8）。

需要注意的是，问题（8.7）不仅涉及求解给定 $s(\pi)$ 下的最优 a 水平，而且还给出了求解最优的合同 $s(\pi)$ 的问题，即一阶条件（8.10）是描绘最优合约 $s(\pi)$ 的条件，而一阶条件（8.8）是刻画给定最优合约 $s(\pi)$ 下的最优的 a 的条件。

一阶条件是一个典型的帕累托最优条件，即努力的期望边际收益等于期望边际成本。

我们称由式（8.8）决定的 a^* 为帕累托最优努力程度。根据一阶条件（8.10），$\lambda=\frac{v'[\pi-s^*(\pi)]}{u'[s^*(\pi)]}$，其中 $s^*(\pi)$ 是最优合约。这说明，在最优合约 $s^*(\pi)$ 处，$\frac{v'}{u'}$ 是一个与 π 无关的常数，因而此时 $\frac{v'}{u'}$ 也是一个与 θ 无关的常数。这时一阶条件（8.8）变为

$$E\left[v'\frac{\partial \pi}{\partial a}\right]=\frac{v'}{u'}\frac{\partial c(a)}{\partial a} \qquad (8.11)$$

8.3.3 最优风险分摊的帕累托最优条件

下面，我们来说明二个一阶条件的经济含义。下面的分析将表明，当我们考虑委托人和代理人在确定情形下的效用函数时，我们就可对一阶条件有一个十分“经济学”的理解。

记 $\bar{v}$ 和 $\bar{u}$ 分别为委托人和代理人在给定 a 下在不确定收入情形下的期望效用函数，则

$$\bar{v}=\int v[\pi-s(\pi)]g(\theta)d\theta$$

$$\bar{u}=\int u[s(\pi)]g(\theta)d\theta$$

如果记 $y(\theta)=\pi(a,\theta)-s[\pi(a,\theta)]$，$z(\theta)=s[\pi(a,\theta)]$，则 $\bar{v}=\int v[y(\theta)]g(\theta)d\theta$，$\bar{u}=\int u[z(\theta)]g(\theta)d\theta$。

当我们不考虑 $y(\theta)$ 和 $z(\theta)$ 是经过具体的产出 π 与 θ 建立联系，而在一般意义上考察

$y(\theta)$ 和 $z(\theta)$ 与 θ 的联系的情形，对于委托人和代理人来说，$y(\theta_1)$ 与 $y(\theta_2)$ 的边际替代率 $\frac{dy(\theta_2)}{dy(\theta_1)}$，以及 $z(\theta_1)$ 对 $z(\theta_2)$ 的边际替代率 $\frac{dz(\theta_2)}{dz(\theta_1)}$ 可分别计算如下：

$$\frac{\partial\overline{v}}{\partial y(\theta_1)}+\frac{\partial\overline{v}}{\partial y(\theta_2)}\frac{dy(\theta_2)}{dy(\theta_1)}=0$$

$$\frac{dy(\theta_2)}{dy(\theta_1)}=-\frac{\dfrac{\partial\overline{v}}{\partial y(\theta_1)}}{\dfrac{\partial\overline{v}}{\partial y(\theta_2)}}=-\frac{v'[y(\theta_1)]g(\theta_1)}{v'[y(\theta_2)]g(\theta_2)}$$

同理有

$$\frac{dz(\theta_2)}{dz(\theta_1)}=-\frac{u'[z(\theta_1)]g(\theta_1)}{u'[z(\theta_2)]g(\theta_2)}$$

在最优合约下，若 $y(\theta)=\pi-s^*(\pi)$， $z(\theta)=s^*(\pi)$，由一阶条件（8.10）有：

$$\lambda=\frac{v'[\pi-s^*(\pi)]}{u'[s^*(\pi)]} \qquad (8.12)$$

$$=\frac{v'[y(\theta)]}{u'[z(\theta)]}$$

=与 θ 无关的常数

于是有 $\frac{v'[g(\theta_1)]}{u'[z(\theta_1)]}=\frac{v'[g(\theta_2)]}{u'[z(\theta_2)]}$

即

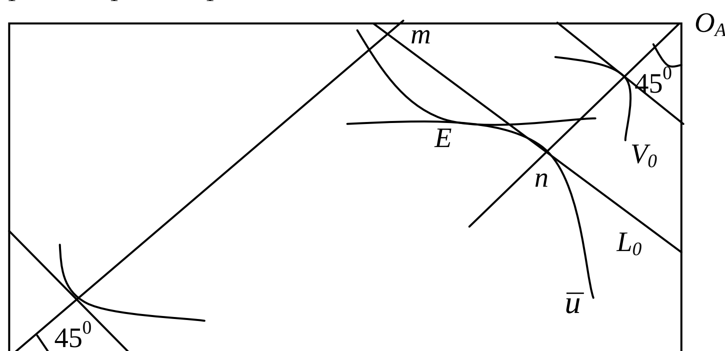
$$-\frac{v'[y(\theta_1)]}{v'[y(\theta_2)]}=-\frac{u'[z(\theta_1)]}{u'[z(\theta_2)]}$$

故有

$$\frac{dy(\theta_2)}{dy(\theta_1)}=\frac{dz(\theta_2)}{dz(\theta_1)}$$

这说明，在最优合约下，在由任意的两种随机情形 θ_1 和 θ_2 下的收入构成的 edgeworth 盒中，委托人与代理人的无差异曲线是相切的，[8]见图 8.1。

$\pi(a,\theta_1)=y(\theta_1)+z(\theta_1)$



O_P $\pi(a,\theta_2)=y(\theta_2)+z(\theta_2)$

图 8.1　帕累托最优风险分摊合约

这是典型的帕累托最优条件。我们来分析，这种帕累托最优是针对什么样的分配来说的。

当一个消费者的收入是在不确定性环境下取得的时，他的效用会因“风险”的引入而与确定性收入环境下有所不同。譬如，当一个消费者不喜欢风险时，他的效用会在给定收入水平下随风险的上升而下降。此时，他宁愿支付一个价格去减少风险，这种购买行为称为“保险”。但是，不是所有消费者在所有收入水平下都是不喜欢风险的。有时，消费者在风险上升时效用会上升。譬如，当购买一张电脑彩票中奖的概率为百万分之一，头奖为一百万元且一张电脑彩票的价格为 2 元钱时，如果潜在的奖励只有百万分之一的一百万元奖金，则买一张电脑彩票的期望收入为$10^6\times10^{-6}=1$元，期望净收入$=1-2=-1$元。但是，即使在这种净期望收入为负数的情况下，也不排除有人会购买电脑彩票。这是因为，即使购买一张彩票的期望收入小于购买彩票的价格，但未来收入是随机的，当未来收入偏离期望收入时，可能中大奖而仅花 2 元钱就获得一百万元。这种可能性会诱使一些“喜欢风险”的人去购买彩票。

这样，消费者在风险面前实际上是有着不同的偏好的。假若我们用k表示风险，且k愈大风险愈大，用w表示期望收入水平，则消费者的效用函数可以表达为

$u=u(w,k)$

并且有$\frac{\partial u}{\partial w}\geq 0$，但不必有$\frac{\partial u}{\partial k}\leq 0$。按照$\frac{\partial u}{\partial k}$的符号，我们可以将消费者对风险的偏好分为三类，且它们都是外生的。

为了解说这三类情形，我们不妨设消费者的效用函数为$u=u(w)$，w为消费者的未来收入。在存在收入上的不确定性时，不妨设收入可能取的数值只有w_1和w_2两种可能，各自发生的概率分别为p_1和p_2。

期望收入为$\overline{w}=p_1w_1+p_2w_2$，$p_1+p_2=1$

V-N-M 期望效用函数为$\overline{u}=p_1u(w_1)+p_2u(w_2)$

1．当消费者不喜欢风险即收入上的不确定性时，他宁愿选择获得一个确定性的收入水平$w=\overline{w}$，也不喜欢在不确定性下（尽管也获得同样的期望收入水平$\overline{w}$，但实际上获得的收入是不确定的）按同样期望收入随机地获得w_1或w_2。这时应有

$$u(\overline{w}) = u(p_1 w_1 + p_2 w_2) \geq p_1 u(w_1) + p_2 u(w_2) = \overline{u}$$

图 8.2 表明：此时要求$u(w)$是凹的，即收入的边际效用递减。

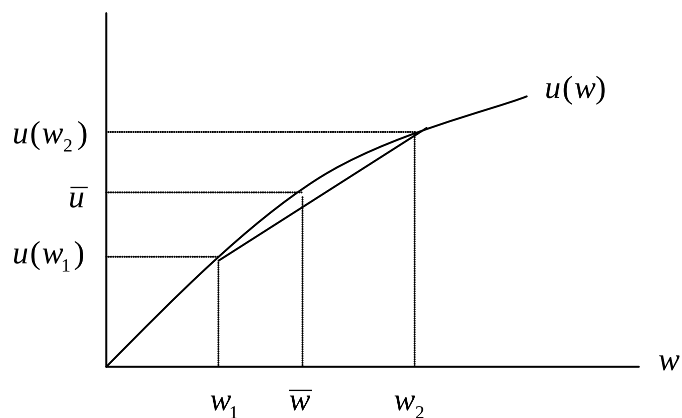


图 8.2　风险规避

称消费者是风险规避（risk-averse）型的，此时有$u'' < 0$。

2．当消费者不介意风险，有

$$u(p_1 w_1 + p_2 w_2) = p_1 u(w_1) + p_2 u(w_2)$$

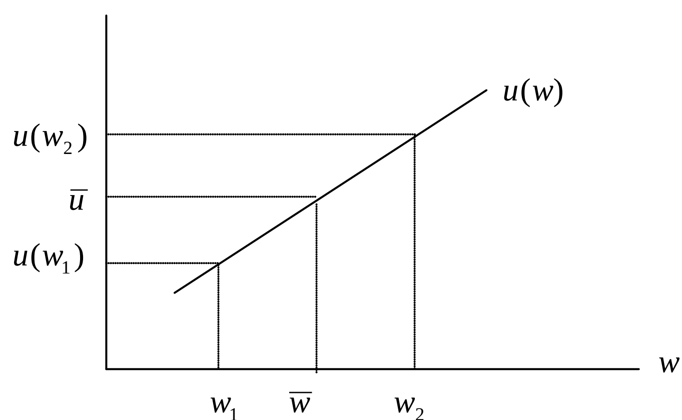


图 8.3　风险中性

称消费者为风险中性（risk-neutral）型的，此时有$u'' = 0$，见图 8.3。

3．当消费者偏好风险时，有

$$u(p_1 w_1 + p_2 w_2) < p_1 u(w_1) + p_2 u(w_2)$$

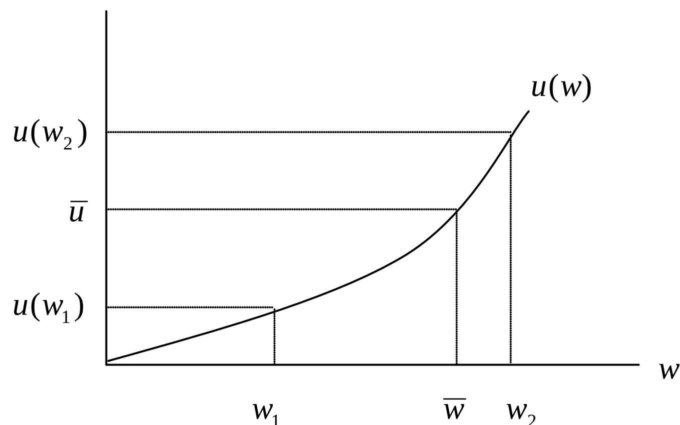


图 8.4　风险偏好

称消费者为风险偏好（risk prefer）型的，此时有$u''>0$。

需要指出的是，在经济学中，尽管有戈森第一定律假定消费者对物品的消费呈边际效用递减趋势，但经济学自身的逻辑体系并不能保证消费者对收入的边际效用也一定呈递减趋势，因而消费者不一定是风险规避的。

显然，在金融市场上进行投机的人一般都是风险偏好者，就象前面提到的电脑彩票购买者是风险偏好者一样，他们希望风险愈大愈好，这样他们的未来收入就愈可能偏离期望收入而使他们获得更大的获取暴利的机会。

一般地，我们可以假定同一个消费者在不同收入水平的变动范围内具有不同的风险态度，如收入较低时是风险规避的，很怕风险，但收入较高后就喜欢风险了，变成了风险偏好者。由此获得的推论是，低收入者投资于债券等低风险金融产品，高收入者投资于股票等高风险金融产品。当然，也不尽然，如在香港，我们发现低收入者是主要的赌马者，而赌马是风险的。一个自然而然产生出来的问题是，有没有一种可测量的指标可用于对消费者对风险的态度进行计量？

显然，当$u''<0$且$|u''|$愈大时，消费者就可被理解为愈怕风险，这是否意味着可用$|u''|=u-u''$作为计量消费者风险规避程度的指标呢？不能，这是因为，效用函数不是唯一的，u在严格单调递增函数的复合下仍为效用函数，若用$-u''$作为刻画消费者规避风险程度的指标，则其数值大小在不同的效用函数下会取不同的数值，有的效用函数给出的$-u''$较大，而另一些效用函数给出的$-u''$较小，是不确定的。那么，如何才能刻画消费者对风险的态度呢？我们希望找到一种指标，它在效用函数的严格递增函数的复合下是一个不变量。但可惜这是做不到的。现在只有后退一步，是否能找到一种指标，它在效用函数的正的严格单调递增线性变换下是不变量。

设 $v = a + bu$， a, b 为常数， $b > 0$

因为 $v'' = bu''$

自然想到构造 $\frac{v''}{v'} = \frac{bu''}{bu'} = \frac{u''}{u'}$

所以，可以用 $r = -\frac{v''}{v'} = -\frac{u''}{u'} = \frac{d\ln\left(\frac{1}{u'}\right)}{dw}$ 作为这种指标，当消费者是风险规避者时，有 $r \geq 0$。

我们称 $r = -\frac{u''}{u'}$ 为消费者的阿罗——普拉特绝对风险规避度（Arrow-Pratt measure of absolute risk aversion）。r 愈大，消费者愈怕风险。

在图 8.1 中，委托人的无差异曲线以 O_P 为原点，代理人的无差异曲线以 O_A 为原点。在 45^0 线上，对应局中人在不同的随机状态下获得相等的收入。这里，我们仅考察 θ 在 θ_1 和 θ_2 之间取值，但 θ_1 和 θ_2 是任意的。这样做的唯一目的是使我们能够在二维图形上进行分析。因此，在 45^0 线上，局中人没有承担任何风险，即 45^0 线是零风险的确定性收入曲线。

下面，我们来计算无差异曲线在对应的确定性收入线上的斜率。对于委托人，有

$$y(\theta_2) = \pi(a,\theta_2) - s[\pi(a,\theta_2)] = \pi(a,\theta_1) - s[\pi(a,\theta_1)] = y(\theta_1)$$

无差异曲线斜率 $= \frac{dy(\theta_1)}{dy(\theta_2)}$（对应于图 8.1 中的坐标轴）

$$= -\frac{v'[y(\theta_2)]g(\theta_2)}{v'[y(\theta_1)]g(\theta_1)}$$

$$= -\frac{g(\theta_2)}{g(\theta_1)}$$

$$= -\frac{f[\pi(a,\theta_2),a]}{f[\pi(a,\theta_1),a]}$$

显然，无差异曲线在对应的确定性收入线上的斜率是相同的。对于代理人而言，也是类似的，并且，在各自的确定性收入线上，两人的无差异曲线的斜率相等。

最优合约的一阶条件（8.10）的几何含义就是：代理人的无差异曲线 $\bar{u}$ 与委托人的无差异曲线 v^0 的切点 E 决定了最优合约在 θ_1 和 θ_2 两种状态下对产出 π 在委托人和代理人之间的分配。

在给定 $a = a^*$ 的情况下，任何 θ 的实现值都给出了一个对应的产出 $\pi(a,\theta)$，而对 $\pi(a,\theta)$ 的分配当 θ 取任意值时实际上就完全决定了合约 $s^*(\pi)$。

下面再计算委托人无差异曲线斜率的系数：

$$\frac{d\left(\frac{dy(\theta_1)}{dy(\theta_2)}\right)}{dy(\theta_2)}=\frac{d\left[-\frac{v'[y(\theta_2)]g(\theta_2)}{v'[y(\theta_1)]g(\theta_1)}\right]}{dy(\theta_2)}$$

$$=-\frac{g(\theta_2)}{g(\theta_1)}\frac{v''[y(\theta_2)]v'[y(\theta_1)]-v'[y(\theta_2)]v''[y(\theta_1)]\frac{dy(\theta_1)}{dy(\theta_2)}}{v'[y(\theta_1)]^2}$$

$$=\frac{g(\theta_2)}{g(\theta_1)}\frac{(-1)}{v'[y(\theta_1)]^2}\Big\{v''[y(\theta_2)]v'[y(\theta_1)]+v'[y(\theta_2)]v''[y(\theta_1)]$$

$$\frac{v'[y(\theta_2)]g(\theta_2}{v'[y(\theta_1)]g(\theta_1)}\Big\}$$

$$=\frac{g(\theta_2)}{g(\theta_1)}\left\{-v''[y(\theta_2)]v'[y(\theta_1)]-v''[y(\theta_1)]\frac{[v'(y(\theta_2))]^2}{v'[y(\theta_1)]}\frac{g(\theta_2)}{g(\theta_1)}\right\}$$

我们假定委托人是风险规避的，则有$v''\le 0$，于是有$\frac{d\left(\frac{dy(\theta_1)}{dy(\theta_2)}\right)}{dy(\theta_2)}\ge 0$。

此时，$\frac{dy(\theta_1)}{dy(\theta_2)}$随$y(\theta_2)$的增加而递增，委托人的无差异曲线凸向原点$O_P$。

类似地，代理人的无差异曲线也凸向原点O_A，见图 8.1。

当$\pi(a,\theta_1)\ne\pi(a,\theta_2)$时，过$O_P$和$O_{A'}$的两条 45^0 线是分离的，此时无差异曲线的切点必然位于两条 45^0 线之间（见图 8.1），这是因为，如果E在某条 45^0 线上，则它也必然在另一条 45^0 线上，这意味着两条 45^0 线重合，这与假定相矛盾；另外，由图 8.1 中的几何关系不难看出，E不会在两条 45^0 线之间的区域以外的其它任何区域。

既然E不在任何一条确定性收入线（45^0 线）上，则每一方的收入都存在不确定性。这就是说，最优合约要求每一方都承担一定的风险。让我们回忆一下本章 8.2.2 中谈到的张五章的工作，这正好证明了张五常关于分成制合约合理性的理论分析。

若委托人是风险中性的，代理人是严格风险规避者，则$v''=0$，$u''<0$，此时$v'(y)$为常数。

由$\frac{dy(\theta_1)}{dy(\theta_2)}=-\frac{g(\theta_2)}{g(\theta_1)}\frac{v'[g(\theta_2)]}{v'[g(\theta_1)]}$

$$=(\text{常数})-\frac{g(\theta_2)}{g(\theta_1)}$$

故委托人的无差异曲线是直线（图 8.1 中的L_0），最优合约在n点。此时代理人的收入是确定的（n在其确定性收入线上），他不承担任何风险。此时所有风险都由委托人承担。这说明，不惧风险者应承担较多的风险。

事实上，此时$v'[g(\theta)]$为常数，则一阶条件（8.10）变为

$$\frac{常数}{u'[s^*(\pi)]}=常数$$

因u'是递减函数，故必有$s^*(\pi)=常数s$，所以，代理人收入与产出π无关。

类似地，若委托人是严格风险规避者，代理人是风险中性者，则$v''<0$，$u''=0$（u'为常数）。

此时代理人的无差异曲线为直线L_0（只要无差异曲线为直线，斜率总为$-\frac{g(\theta_2)}{g(\theta_1)}$），最优合约在$m$点，此时委托人得到固定收入$y(\theta_1)=y(\theta_2)=\pi-s(\pi)=y_0$，代理人承担全部风险$(z=s(\pi)=\pi-y_0)$。

若委托人和代理人都是风险中性者$(v''=u''=0)$，则委托人和代理人的无差异曲线都是直线，且两人的无差异曲线有相同斜率$-\frac{g(\theta_2)}{g(\theta_1)}$。这时，直线$L_0$上的所有点都是最优的。

一般地，一阶条件$\frac{v'[\pi-s^*(\pi)]}{u'[s^*(\pi)]}=\lambda$隐含地定义了最优合约$s^*(\pi)$。

我们在一阶条件两端对π求导，得：

$$\frac{v''[\pi-s^*(\pi)]\left(1-\frac{ds^*}{d\pi}\right)}{u'[s^*(\pi)]}-\frac{v'[\pi-s^*(\pi)]u''[s^*(\pi)]}{u'[s^*(\pi)]^2}\left(\frac{ds^*}{d\pi}\right)=0$$

故$v''[\pi-s^*(\pi)]\left(1-\frac{ds^*}{d\pi}\right)u'[s^*(\pi)]-v'[\pi-s^*(\pi)]u''[s^*(\pi)]\frac{ds^*}{d\pi}=0$

两边同除u'，得

$$-v''\left(1-\frac{ds^*}{d\pi}\right)-\frac{v'[\pi-s^*(\pi)]}{u'[s^*(\pi)]}u''[s^*(\pi)]\frac{ds^*}{d\pi}=0$$

由一阶条件$\frac{v'[\pi-s^*(\pi)]}{u'[s^*(\pi)]}=\lambda$，代入上式得：

$$-v''\left(1-\frac{ds^*}{d\pi}\right)+\lambda u''\frac{ds^*}{d\pi}=0$$

令$\rho_p=-\frac{v''}{v'}$，$\rho_A=-\frac{u''}{u'}$分别是委托人和代理人的绝对风险规避度，解上式得：

$$\frac{ds^*}{d\pi}=\frac{v''}{v''+\frac{v'}{u'}u''} \qquad （8,13）$$

$$=\frac{-\frac{v''}{v'}}{-\frac{v''}{v'}-\frac{u''}{u'}}$$

$$= \frac{\rho_p}{\rho_p + \rho_A}$$

式（8.13）说明：代理人的支付 s^* 与产出 π 的关系完全由绝对风险规避度的比率决定（即 $\frac{ds^*}{d\pi} = \frac{1}{1+\frac{\rho_p}{\rho_A}}$ ）。

给定 $\rho_p > 0$ ， $\rho_A > 0$ （即双方均为风险规避者），则代理人的支付 s^* 是 π 的增函数（ $\frac{ds^*}{d\pi} > 0$ ），但上升幅度小于 π 上升的幅度（即 $\frac{ds^*}{d\pi} < 1$ ）。当 $\rho_p = 0$ 时， $\frac{ds^*}{d\pi} = 0$ ，此时 s^* 与 π 无关；

当 $\rho_A = 0$ 时， $\frac{ds^*}{d\pi} = 1$ ， s^* 的增幅与 π 相同。

当委托人和代理人都具有不变的绝对风险规避度时[9]，即若 ρ_p 和 ρ_A 与各自的收入水平无关时（都为常数），则对式（8.13）积分：

$$s^* = \alpha + \beta\pi \qquad (8.14)$$

此时，最优合约 $s^*(\pi)$ 是线性的，其中 $\beta = \frac{\rho_p}{\rho_p + \rho_A}$ ， α 为积分常数。

我们以上的分析是在给定最优努力水平 $a = a^*$ 下的情形进行的。此时，影响产出 π 的因素只有随机因素 θ 。在给定 a 的情形，对收入的分配不仅有存在参与约束给出的约束，而且还存在风险分摊的帕累托最优原则给出的约束。所谓风险分摊的帕累托最优原则，就是指当我们将风险引入效用函数时，不怕风险（如风险中性者）应承担全部风险，或绝对风险规避度较小的风险规避者应承担较大的风险。在委托人与代理人的无差异曲线的切点处，就是这种风险分摊的帕累托最优点或帕累托最优风险分摊点。

一阶条件（8.10）说明，当信息对称时，帕累托最优风险分摊是能够实现的。

8.3.4 最优努力水平的决定：激励问题

下面，我们再来看决定最优努力程度 a^* 的一阶条件。最优努力程度的刻画是激励问题，其一阶条件就是式（8.8）。

设委托人是风险中性的，则 $v'' = 0$ ， v' 为常数（委托人是公司时，可假定是风险中性的，见 Mereon（19　）。

不妨设 $v' = 1$ ，则一阶条件（8.8）变为：

$$E\frac{\partial \pi}{\partial a} = \left(E\frac{1}{u'}\right)\frac{\partial c}{\partial a}$$

据前面的分析，一阶条件（8.12）保证在最优风险分摊下，代理人的收入是确定的

（给定每一个 a ）。故

$u'[s^*(\pi)] = u'(s^0)$ 是确定的，于是有

$$E\frac{1}{u'} = \frac{1}{u}$$

$$E\frac{\partial \pi}{\partial a} = \frac{1}{u'}\frac{\partial c}{\partial a} \tag{8.15}$$

其中，$E\frac{\partial \pi}{\partial a} = \int \frac{\partial \pi(a,\theta)}{\partial a} g(\theta) d\theta$ 是期望边际产出，$\frac{1}{u'}\frac{\partial \pi}{\partial a}$ 是代理人在货币收入与努力之间的边际替代率。为了理解后一点，只需注意到代理人的净效用函数为 $U = u[s(\pi)] - c(a)$，边际替代率为

$$\left|\frac{\frac{\partial U}{\partial a}}{\frac{\partial U}{\partial s}}\right| = \left|\frac{-\frac{\partial c}{\partial a}}{\frac{\partial u}{\partial s}}\right|$$

$$= \frac{1}{u'}\frac{\partial c}{\partial a}$$

因为代理人的净效用函数为：

$$U = u[s(\pi)] - c(a)$$

无差异曲线为方程 $U_0 = u[s(\pi)] - c(a)$ 所隐含的曲线，其中 U_0 为常数。

在方程两端求微分：

$$0 = u'[s(\pi)]ds - \frac{\partial c}{\partial a}da$$

故 $\frac{ds}{da} = \frac{\frac{\partial c}{\partial a}}{u'[s(\pi)]} = \frac{c'}{u'}$，这是无差异曲线的斜率

于是 $\frac{d^2 s}{da^2} = \frac{c''u' - c'u''\frac{ds}{da}}{u'^2}$

$$= \frac{c''u' - c'u''\frac{c'}{u'}}{u'^2}$$

$$= \frac{c''u' - u''\frac{c'^2}{u'}}{u'^2} > 0$$

所以，无差异曲线是凸的，见图 8.5。

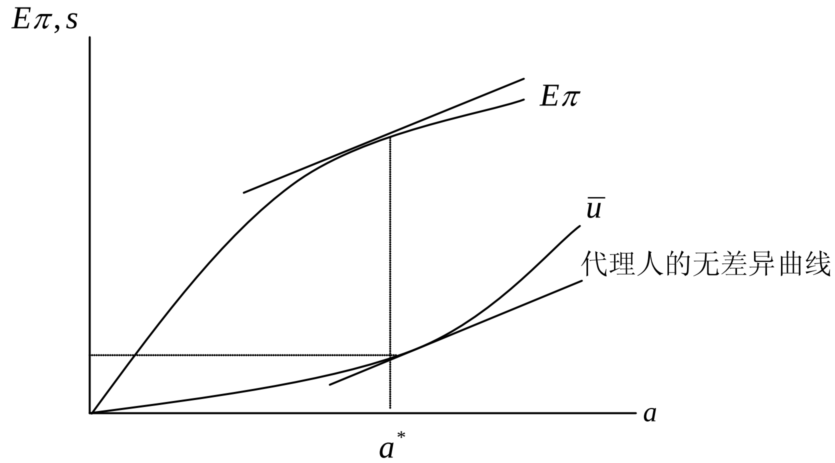


图 8.5　最优努力水平

因有 $E\pi(a,\theta)=\int \pi(a,\theta)g(\theta)d\theta$

故 $\frac{\partial E\pi}{\partial a}=\int \frac{\partial \pi(a,\theta)}{\partial a}g(\theta)d\theta$

据前面的假定：$\pi(a,\theta)$ 在给定的 θ 下是 a 的严格递增凹函数。

所以，$\frac{\partial \pi}{\partial a}>0$

$\frac{\partial E\pi}{\partial a}>0$

故 $E\pi$ 是 a 的严格递增函数。

又因为 $\pi(a,\theta)$ 是 a 的严格凹函数，于是

$\frac{\partial^2 \pi}{\partial a^2}<0$

或 $\frac{\partial^2 E\pi}{\partial a^2}=\int \frac{\partial^2 \pi(a,\theta)}{\partial a^2}g(\theta)d\theta<0$

所以，$E\pi$ 是 a 的严格凹函数，见图 8.5。

式（8.5）说明 $E\pi$ 线的斜率 $\frac{\partial E\pi}{\partial a}$ 等于代理人的无差异曲线斜率 $\frac{\frac{\partial c}{\partial a}}{u'}$。

在 a^* 点处，期望边际产出等于边际替代率，故 a^* 是最优努力水平。

因为这里假定委托人是风险中性的，根据最优风险分摊的条件，s^* 是独立于 θ 的（代理人的收入是确定的 s^*）。

显然，委托人最优的支付水平 s^* 满足

$\int u(s^*)g(\theta)d\theta-c(a^*)=\bar{u}$

因 s^* 是与 θ 无关的，故 $\bar{u}=u(s^*)-c(a^*)$

最优支付 s^* 满足：$u(s^*)=\bar{u}+c(a^*)$

类似地，当代理人是风险中性的时，有

$u''=0$，u' 为常数，设为 1。

由前面的分析知此时最优风险分摊意味着委托人获得的收入是确定的 y^0，使得 v' 不变。

由一阶条件（8.8）：

$$v'E\left[\frac{\partial\pi}{\partial a}-\frac{1}{u'}\frac{\partial c}{\partial a}\right]=0$$

故 $v'E\dfrac{\partial\pi}{\partial a}=E\dfrac{v'}{u'}\dfrac{\partial c}{\partial a}$

即努力的边际收益等于努力的边际成本，此时代理人的收入为 $\pi(a^*,\theta)-y^0$。

最后，我们还需说明委托人设计什么样的机制使代理人选择行动 a^*。这是一个博弈。

委托人可以设计如下的强制性合约：

$$s(\pi)=\begin{cases}s^*(\pi)=s^*[\pi(a^*,\theta)], a\geq a^*\\ \underline{s} \qquad\qquad\qquad\qquad a<a^*\end{cases}$$

其中，$\underline{s}$ 是充分小的正数。

即委托人要求代理人选择 a^*，若预测到代理人选择了 $a\geq a^*$，委托人根据 $s^*[\pi(a^*,\theta)]$。

支付给代理人；否则，代理人得到 $\underline{s}$。只要 $\underline{s}$ 足够小，代理人就不会选择 $a<a^*$，因为代理人在选择 $a<a^*$ 时的净效用为

$$\int u[\underline{s}]f(\pi,a)d\pi-c(a)$$

$$=a(\underline{s})-c(a)$$

只要 $\underline{s}$ 足够小，代理人选较小的 a 会带来负的净效用，还不如选 a^* 获得保留效用 $\bar{u}>0$。故只需有

$u(\underline{s})=\bar{u}$

此时，$u(\underline{s})-c(a)<\bar{u}$

因为 $c'(a)>0$，代理人实际上不会选 $a>a^*$。博弈均衡结果是代理人选择 $a=a^*$。

本小节的基本结论是，当委托人可以无成本地观测到代理人的努力水平 a 时，风险的最优分摊和激励问题可以独立解决，帕累托最优风险分摊和帕累托最优努力水平可同时实现。

8.4 最优合约：不对称信息的情形

8.4.1 道德风险

在上一节里，我们假定委托人可以无成本地观测到代理人的行为，这是个不现实的假定。在接下来的这一节里。我们将放弃这一不现实的假定，而用委托人不能无成本地观测到代理人的行为的不对称信息假定来取代它。我们的问题是，在这种情况下，委托人如何设计合约来诱使代理人选择最有利于委托人利益的行动？

我们将证明，除非代理人是风险中性的，此时帕累托最优的努力水平 a^* 是无法实现的。

这时，给定合约 $s^*[\pi(a^*,\theta)]$，代理人将由下述的激励相容约束来决定选择行动 a：

$$\max_a \int u\{s^*[\pi(a^*,\theta)]\}g(\theta)d\theta - c(a)$$

一阶条件

$$\int u' \frac{\partial s^*}{\partial \pi}\frac{\partial \pi}{\partial a} g(\theta)d\theta - \frac{\partial c}{\partial a} = 0$$

即

$$E\left[u' \frac{\partial s^*}{\partial \pi}\frac{\partial \pi}{\partial a} - \frac{\partial c}{\partial a}\right] = 0$$

$$Eu'\left[\frac{\partial s^*}{\partial \pi}\frac{\partial \pi}{\partial a} - \frac{1}{u'}\frac{\partial c}{\partial a}\right] = 0 \tag{8.16}$$

今设式（8.16）的解为 a^+，即 a^+ 是激励相容约束解，它是代理人选择的努力水平。

由于激励相容约束（8.16）与参与约束最大化问题（8.8）是不同的条件，故它们的解一般不会相等，即一般有 $a^+ \neq a^*$。

事实上，我们可证明有 $a^+ < a^*$。

根据式（8.13）$\dfrac{ds^*}{d\pi} = \dfrac{\rho_p}{\rho_p + \rho_A}$

故一般有 $\dfrac{\partial s^*}{\partial \pi} < 1$。

又由 $c'' > 0$ 的假定，c' 是递增的，将 $u' = \dfrac{v'}{\lambda}$ 代入式（8.16）

$$Ev'\left[\frac{\partial s^*}{\partial \pi}\frac{\partial \pi}{\partial a}\right] = \lambda \frac{\partial c}{\partial a}$$

若没有 $a^+ < a^*$，则有 $a^+ \geq a^*$。

此时

$$Ev'\left[\frac{\partial s^*}{\partial \pi}\frac{\partial \pi}{\partial a}\right]\Bigg|_{a^+} = \lambda\frac{\partial c}{\partial a}\Bigg|_{a^+} \geq \lambda\frac{\partial c}{\partial a}\Bigg|_{a^*}$$

（λ是与a,π无关的常数，这是拉格朗日乘数的性质）。

据式（8.8），上式又等于$Ev'\frac{\partial \pi}{\partial a}\Big|_{a^*}$，故有$Ev'\left[\frac{\partial s^*}{\partial \pi}\frac{\partial \pi}{\partial a}\right]\Big|_{a^+} \geq Ev'\frac{\partial \pi}{\partial a}\Big|_{a^*}$。

又因为$v'\frac{\partial s^*}{\partial \pi}\frac{\partial \pi}{\partial a}\Big|_{a^+} < v'\frac{\partial \pi}{\partial a}\Big|_{a^*}$在任何$\pi$下都成立（设$\frac{\partial s^*}{\partial \pi}<1$），因$\frac{\partial \pi}{\partial a}$是$a$的减函数，即$\pi$是$a$的凹函数。

又由于$\frac{d[\pi - s^*(\pi)]}{d\pi} = 1 - \frac{\partial s^*}{\partial \pi} = \frac{\rho_A}{\rho_p + \rho_A} \geq 0$

以及π是a的增函数和v'是减函数的假定，故

$v'\big|_{a^+} \leq v'\big|_{a^*}$

于是$Ev'\left[\frac{\partial s^*}{\partial \pi}\frac{\partial \pi}{\partial a}\right]\Big|_{a^+} < Ev'\frac{\partial \pi}{\partial a}\Big|_{a^*}$。

这导致矛盾。

所以，$a^+ < a^*$，即代理人选择的努力水平a^+小于帕累托最优努力水平a^*。在上面的分析中，我们用“$|_{a^+}$”和“$|_{a^*}$”分别表示a在a^+和a^*处取值。

譬如，若委托人是风险中性的，且假定$c'(0)=0$，则帕累托最优风险分摊要求有$s^*(\pi)=$常数s_0。

因为式（8.8）给出的$a>0$（因$E\frac{\partial \pi}{\partial a}>0$）但由式（8.16）给出（因$\frac{\partial s^*}{\partial \pi}=0$）：$\frac{\partial c}{\partial a}=0$，它当且仅当$a^+=0$成立。

即若收入与工作努力程度无关，代理人为什么要努力工作呢？从直观上看，给定$s^*(\pi)$，则对委托人来说是最优的努力水平a^*（由式（8.8）给出），对代理人来说却并不是最优的，因为对代理人来说最优的a由式（8.16）决定。

因此，倘若委托人不能观测到a，则代理人会作如下的选择：私下选择$a<a^*$（按式（8.16）选择）以改进自己的福利水平。因为利润水平π不仅对代理人的努力水平有关，而且还受外生变量θ的影响，代理人可将低利润的出现归咎于不利的θ的影响，从而逃避委托人的职责。

因为委托人不能观测a，自然也就不能证明低利润是代理人没有努力工作的结果。

经济学家称代理人的这种行为为“道德风险”（moral hazard）。

8.4.2 不存在道德风险的情形

我们将看到，当委托人不能观测代理人的工作努力水平时，最优激励合约让代理人承担比对称信息下更大的风险，从而导致帕累托最优风险分摊不能实现。

但是，即使代理人的行动不可观测，但若代理人是风险中性的，帕累托最优努力水平和最优风险分摊也可同时实现，不会出现道德风险问题。

事实上，由式（8.13)，此时有 $\frac{\partial s^*}{\partial \pi}=1$，即 $s^*=\pi-y_0$，y_0 是委托人的确定性收入（由最优风险分摊条件）。

故 v' 和 u' 都是常数，设 $u'=v'=1$，因为有 $v'=v'(y_0)$，式（8.8）变为

$$E\frac{\partial \pi}{\partial a}=\frac{\partial c}{\partial a}$$

式（8.16）变为 $E\frac{\partial \pi}{\partial a}=\frac{\partial c}{\partial a}$

两式相同，故 $a^+=a^*$。

从直观上看，当代理人承担全部风险时，没有外部效应，代理人就如同为自己工作一样，不会有偷懒的动机。且因代理人是风险中性的时候，委托人的保险成本为零，即委托人将风险加于代理人且由此用于激励代理人时并未支付成本，不存在保险与激励之间的矛盾。

8.5 不对称信息下的最优合约：模型分析

在本节，我们将对决定不对称信息下的最优合约进行模型分析。本节将使用分布函数的参数化方法。本节将分两种情形展开讨论，即先对一个简化的模型展开讨论，然后给出一般化模型。

8.5.1 简化的模型

以下假定是本小节中分析的基本出发点：

（1） a 只取 L(Lazy)和 H(work hard)两个数值；

（2） $\pi\in[\underline{\pi},\overline{\pi}]$；

（3）当 $a=H$， π 的分布函数和分布密度分别为 $F_H(\pi)$ 和 $f_H(\pi)$；

当 $a=L$， π 的分布函数和分布密度分别为 $F_L(\pi)$ 和 $f_L(\pi)$。

前面已假定 $\pi(a,\theta)$ 是 a 的增函数，在这里，对于随机变量 π，该假定的一个重新表述为：

分布函数是一阶随机占优（the first-order stochastic dominance)，即：

对于任意的$\pi \in [\underline{\pi}, \overline{\pi}]$，有$F_H(\pi) \le F_L(\pi)$且至少对某些$\pi$严格不等式成立。其意为：勤奋工作时高利润的概率$[1-F_H(\pi)]$大于偷懒时高利润的概率$1-F_L(\pi)$，即$1-F_H(\pi) \ge 1-F_L(\pi)$。

当a是连续变量且$F(\pi,a)$对a可微分时，该假定就是$\frac{\partial F}{\partial a} < 0$。

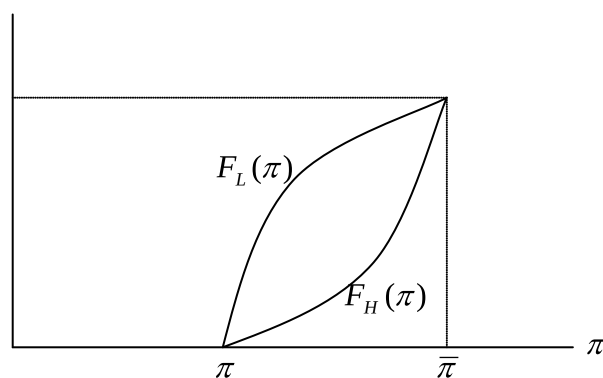


图 8.6　一阶随机占优条件

（4）$c(H) > c(L)$，即勤奋工作的成本比偷懒的成本高。

此时，委托人的最大化问题表述为[10]：

$$\max_{s(\pi)} \int v[\pi - s(\pi)] f_H(\pi) d\pi$$

$$s.t \quad (IR) \quad \int u[s(\pi)] f_H(\pi) d\pi - c(H) \ge \overline{u}$$

$$(IC) \quad \int u[s(\pi)] f_H(\pi) d\pi - c(H) \ge \int u[s(\pi)] f_L(\pi) d\pi - c(L)$$

我们只须考虑解为$a = H$的情形（见注释[10]）。

构造拉格朗日函数

$$\begin{aligned} L[s(\pi),\lambda,u] = &\int v[\pi - s(\pi)] f_H(\pi) d\pi \\ &+ \lambda\left\{\int u[s(\pi)] f_H(\pi) d\pi - c(H) - \overline{u}\right\} \\ &+ \mu\left\{\int \overline{u}[s(\pi)] f_H(\pi) d\pi - c(H) - \int u[s(\pi)] f_L(\pi) d\pi + c(L)\right\} \end{aligned}$$

一阶条件：

$$\frac{\partial L}{\partial s} = \int (-v' f_H) d\pi + \lambda \int u' f_H d\pi + \mu\left[\int u' f_H d\pi - \int u' f_L d\pi\right] = 0$$

$$\int \left[-v' f_H + \lambda u' f_H + \mu u' f_H - u' f_L\right] d\pi = 0$$

$$\int \left[-v' + \lambda u' + \mu u' - \frac{u' f_L}{f_H}\right] f_H d\pi = 0$$

在这里，如果假定委托人知道f_H，则可得到充分条件

$$-v' + \lambda u' + \mu u' - u'\frac{f_L}{f_H} = 0$$

但若假定委托人不知道 f_H，则上式就是充分必要条件。充分条件意味着仅从该条件导出的合约不一定是最优的合约。

此时有

$$\frac{v'[\pi - s(\pi)]}{u'[s(\pi)]} = \lambda + \mu\left(1 - \frac{f_L}{f_H}\right) \qquad (8.17)$$

式（8.17）被称为“Mirrlees—Holmstrom”条件，其中λ，μ为拉格朗日乘数。

在式（8.17）中，若$\mu = 0$，则得到前面已导出的帕累托最优风险分摊条件（8.10），但因$\mu = 0$意味着不考虑激励相容条件，故猜想有$\mu \neq 0$。Holmserom（1979）证明有$\mu > 0$。

这说明：非对称信息情形的最优合约不同于对称信息情形的最优合约。

在式（8.17）中有一个比值$\frac{f_L}{f_H}$，在统计学中被称为似然率（likelihood ratio），其统计学意义为：

假设观测者不知道随机变量π的样本到底来自分布$f_H(a = H)$还是来自分布$f_L(a = L)$，但知道当是来自f_H时，π发生的概率为$f_H\Delta\pi$，当是来自f_L时，π发生的概率为$f_L\Delta\pi$。故似然率$\frac{f_L}{f_H} = \frac{f_L\Delta\pi}{f_H\Delta\pi}$刻画了两种概率之比。在贝叶斯（Bayes）统计学中，用$\frac{f_L}{f_H}$测度观测者对分布的概率判断。

当观测样本提供的$f_H\Delta\pi$的估计为p_H时，提供的$f_L\Delta\pi$的估计为p_L时（如在同样的观测值π下，N次调查（如果可能进行这样的调查的话）发现有$n \le N$次有$a = L$，$N - n$次有$a = H$，则似然比$\frac{n/N}{N - n/N} = \frac{n}{N - n}$可用来判断分布是$f_L$（或$f_H$）的可能性有多大，如$\frac{n}{N - n} = 1$，说明观测者不能得到任何进一步的信息，所以，判断分布是f_H和f_L的可能性是相同的。当$\frac{n}{N - n}$较大时，说明分布为f_L的可能性较大。）

显然，这种样本的估计值应为$\frac{f_L}{f_H}$。

式（.17）告诉我们：代理人的收入$s(\pi)$随似然率$\frac{f_L}{f_H}$而变。此时，代理人的收入比对称信息下具有更大的波动。

譬如，当委托人是风险中性的时，$v' = 1$。在对称信息下，最优风险分摊意味着代理人得到固定收入，不承担任何风险。但在非对称信息下，$v' = 1$并不能保证$s(\pi)$为固

定数（由式（8.17）），故此时代理人也要承担一些风险。这就是由非对称信息导致的激励与保险之间的权衡（trade-off）（因为让代理人承担部分风险有利于激励代理人）。

下面，我们进一步对这一点加以说明。

记 $s_\lambda(\pi)$ 是一阶条件（8.8）决定的最优风险分摊合约（即前述的 $s^*(\pi)$）， $s(\pi)$ 是一阶条件（8.17）给出的激励合约，则由式（8.8）和（8.17）有：

$$\frac{v'[\pi - s_\lambda(\pi)]}{u'[s_\lambda(\pi)]} = \lambda_1$$

$$\frac{v'[\pi - s(\pi)]}{u'[s(\pi)]} = \lambda_2 + \mu\left(1 - \frac{f_L}{f_H}\right)$$

由最优动态控制数学中的庞特里亚金最大值原理可知有 $\lambda_1 = \lambda_2$（它们都是 IR 条件的拉格朗日乘数），已知 $\mu > 0$ 。因为 $\frac{v'(y)}{u'(z)}$ 是 y 的减函数和 z 的增函数，故当 $f_L(\pi) \geq f_H(\pi)$ ，有

$$\frac{v'[\pi - s_\lambda(\pi)]}{u'[s_\lambda(\pi)]} \geq \frac{v'[\pi - s(\pi)]}{u'[s(\pi)]}$$

故必有 $\pi - s_\lambda(\pi) \leq \pi - s(\pi)$

当且仅当 $s_\lambda(\pi) \geq s(\pi)$ 成立。因为否则有 $s_\lambda(\pi) < s(\pi)$ ，得

$\pi - s_\lambda(\pi) \geq \pi - s(\pi)$

所以， $\frac{v'[\pi - s_\lambda(\pi)]}{u'[s_\lambda(\pi)]} < \frac{v'[\pi - s(\pi)]}{u'[s(\pi)]}$

得到 $\begin{cases} s(\pi) \leq s_\lambda(\pi), \text{若} f_L(\pi) \geq f_H(\pi) \\ s(\pi) > s_\lambda(\pi), \text{若} f_L(\pi) < f_H(\pi) \end{cases}$

即：给定产出 π ，若 π 胡代理人选 $a = L$ 时出现的概率大于他选 $a = H$ 出现的概率，则代理人在该利润时的收入所得向下调整；反之亦反。

这反映了似然率 $\frac{f_L}{f_H}$ 包含的信息量。当 $f_L > f_H$ 时，更大可能的是 π 来自 $a = L$ 的分布，故委托人才支付给代理人作为他可能选 L 的处罚。

委托人并不能从观测到的 π 推出任何东西。因为在均衡下，委托人准确地知道代理人选择了什么，尽管他并不能观测到代理人的选择。但为了达到均衡，委托人的战略（或最优激励合约 $s(\pi)$ ）似乎是依赖于这样一种统计推断：委托人根据观测到的 π 推断代理人是选了 L 或 H，进而据此对代理人实行奖惩。譬如，若委托人推断多半有 $a = L$ ，就实行惩罚， $s(\pi) < s_\lambda(\pi)$ ，反之亦反。

但从另一方面看，委托人似乎是在根据贝叶斯法则从观测到的 π 修正他对代理人选 $a = H$ 的后验概率，并据此修正结果对代理人进行奖惩。

我们记$\gamma = \Pr ob(H)$为委托人对代理人选$a = H$的先验概率，$\tilde{\gamma}(\pi) = \Pr ob(H \mid \pi)$为委托人在预测到$\pi$时认为代理人选$a = H$的后验概率。

据贝叶斯法则有

$$\tilde{\gamma}(\pi) = \frac{P(H,\pi)}{P(\pi)}$$

$$= \frac{P(\pi / H)P(H)}{P(\pi / H)P(H) + P(\pi / L)P(L)}$$

$$= \frac{f_H \gamma}{f_H \gamma + f_L (1-\gamma)}$$

这里，我们用“P”表示概率。

故$[\gamma - \gamma\tilde{\gamma}(\pi)] f_H = (1-\gamma) f_L \tilde{\gamma}(\pi)$

所以，$\frac{f_L}{f_H} = \frac{\gamma - \gamma\tilde{\gamma}(\pi)}{\tilde{\gamma}(\pi)(1-\gamma)}$

代入式（8.17）

$$\frac{v'[\pi - s(\pi)]}{u'[s(\pi)]} = \lambda + \mu\left[\frac{\tilde{\gamma}(\pi) - \gamma}{\tilde{\gamma}(\pi)(1-\gamma)}\right] \qquad (8.18)$$

若观测到的π使委托人向下修正了代理人选$a = H$的概率，即$\tilde{\gamma}(\pi) < \gamma$，则

$$\frac{v'[\pi - s(\pi)]}{u'[s(\pi)]} < \lambda$$

即有$s(\pi) < s_\lambda(\pi)$（道理与前面相同）

所以，代理人受到惩罚。反之亦反。

需要指出的是：事实上，π是通过后验概率$\tilde{\gamma}(\pi)$（或者说似然率$\frac{f_L}{f_H}$）影响到代理人的收入$s(\pi)$。如果说s是依赖于π，那么可以说这并非因为π自身的物质价值所致，而是因为它提供的信息量价值。

由此，条件（8.17）对最优合约$s(\pi)$的具体形式可以说是没有任何限制的，即任何形式的$s(\pi)$都可能出现（Hart 与 Holmstrom（1987）对此有专门的讨论）。甚至条件（8.17）给出的$s(\pi)$还可能不是单调的，即较高的π不一定意味着代理得到较高的报酬s。如下例所得：

例 8.1，设当$a = L$时，π有两种可能的取值

$$\pi = \begin{cases} -100 \\ 100 \end{cases}，且\ f_L(-100) = f_L(100) = 0.5$$

当$a = H$时，π也有两种可能的取值。

$$\pi = \begin{cases} -100 \\ 500 \end{cases}，且\ f_H(-100) = f_H(500) = 0.5$$

这种分布函数满足一阶随机占优条件：

对所有的π，有 $F_H(\pi) \le F_L(\pi)$

因为 $F_H(-100) = F_L(-100) = 0.5$

$F_H(100) = 0.5 < F_L(100) = 1$

$F_H(500) = 1 = F_L(500) = 1$

故 $\left.\frac{f_L}{f_H}\right|_{\pi=-100} = \frac{0.5}{0.5} = 1$

$\left.\frac{f_L}{f_H}\right|_{\pi=100} = \frac{0.5}{0} = \infty > \left.\frac{f_L}{f_H}\right|_{\pi=-100}$

据式（8.8），有：$s(100) < s(-100)$

所以，在π较大时代理人收入还应减少，这是因为当$\pi = -100$，委托人认为代理人选L和H的可能性相同（上面给出的概率相同），但当$\pi = 100$时，则委托人可以肯定代理人选了L（因为 $f_H(100) = 0, f_L(100) > 0$），故对其进行惩罚。

从式（8.17）可知，$s(\pi)$是$\frac{f_L}{f_H}$的单调减函数，即$\frac{f_L}{f_H}$愈大，$s(\pi)$就愈小。如果$\frac{f_L}{f_H}$是π的减函数，则$s(\pi)$就是π的增函数。

当π较大时，$\frac{f_L}{f_H}$较小。代理人就有较大的可能选$a = H$，故给予奖励，即$s(\pi)$较大。这符合大多数我们所观测到的现实。

我们假定：$\frac{f_L}{f_H}$是π的减函数。称此假定为：单调似然率性质假定（MLRP, monotone likehood ratio property）。在现实生活中，我们观测到的大多数合约$s(\pi)$都满足$s(\pi)$对π的单调递增性。如企业经理的报酬随企业利润的增加而增加，这一事实可能意味着绝大多数分布函数具有单调似然率性质（或许还有其它的解释，如当对代理人不利的结果出现时（如当$s(\pi)$规定π愈大时，$s(\pi)$愈小），代理人可将“利润”扔掉从而使π变小，s变大，故若$s(\pi)$规定π大时对代理人进行惩罚实际上是不可行的）。在中国，某些国有企业的经理害怕企业赚钱而导致自身位置的不稳定，可能会在盈利的情况下将帐目“做”成“亏损”，因为对这些国有企业来说，较大的π很可能意味着$s(\pi)$小[11]。这可能也是不满足单调似然率假定的例子。

但是，无论如何，$s(\pi)$的单调性并非是一般的性质。图 8.7 给出的分布函数就不具有单调似然率的性质。

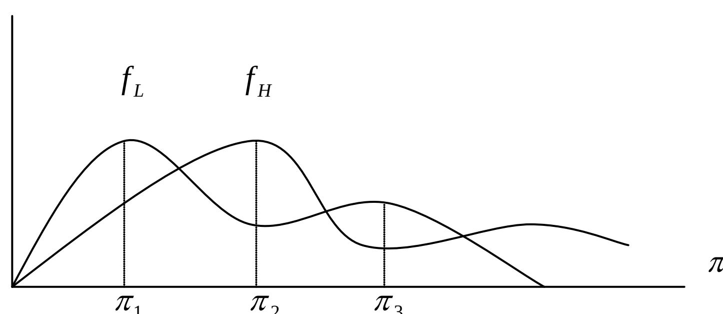


图 8.7　不具有单调似然率性质的分布

显然，有 $\left.\frac{f_L}{f_H}\right|_{\pi_1} > 1 > \left.\frac{f_L}{f_H}\right|_{\pi_2} < \left.\frac{f_L}{f_H}\right|_{\pi_3} > 1$

8.4.2 多观测变量的情形

正如我们在前面所说的，如何设计一个“指标体系”用于合约得以实施的可观测变量组群，是一个需要加以讨论的问题。委托人是根据可观测变量给予代理人行动选择进行奖惩的，但这些可观测变量应该包括哪些变量，在本小节中将给出一个原则性的方法用于判定变量是否可被纳入合约所依赖的可观测变量组群。

在委托—代理理论中，还有一个重要的问题是研究在激励合约中选取什么样的变量作为观测变量可使委托人的效用达到最大化。

假设除 π 之外，委托人还可以不花费成本地观测到另一变量 z 。观测另外变量 z 的目的是节约委托人的风险成本，此时代理人承担较小的风险，但观测可能又有额外的成本。

设 $x=(\pi,z)$ ，如在股东激励经理的场合，π 是利润，z 是与企业运行环境有关的外生变量如货币供给或是另一个企业的利润。当另外的变量可以在较低观测成本下获取关于 a （或 θ ）的更加准确的信息，委托人就毋须给予代理人更多的风险收入（让代理人承担更大风险）从而支付更多的风险成本下对其进行奖惩。这就是观测更多变量的经济学理由。

假定 z 和 a 和（或）θ 有关，即 $z=z(a,\theta)$，从而委托人可从观测 z 中获取关于 θ（或 a ）的更多信息。

现在的问题是：在什么条件下，委托人对代理人的奖惩 s 不仅依赖于 π ，而且还依赖于 z ？即 $s=s(\pi,z)$ ？即 z 可提供关于 a 的更多的信息。

设 $a=L$ 和 H 时 $x=(\pi,z)$ 的联合分布密度分别为 $h_L(\pi,z)$ 和 $h_H(\pi,z)$ ，则当 $s=s(\pi,z)$ 时，委托人的问题是选 $s(\pi,z)$ 解下列最大化问题：

$$\max_{s(\pi,z)} \int_\pi\int_z v[\pi - s(\pi)]h_H(\pi,z)dzd\pi$$

$$s.t \quad (IR) \quad \int_\pi\int_z u[s(\pi)]h_H(\pi,z)dzd\pi - c(H) \geq \bar{u}$$

$$(IC) \quad \int_\pi\int_z u[s(\pi)]h_H(\pi,z)dzd\pi - c(H) \geq \int_\pi\int_z u[s(\pi)]h_L(\pi,z)dzd\pi - c(L)$$

假设 $a = H$ 对委托人来说比 $a = L$ 更有利。

拉格朗日函数为

$$L(s,\lambda,\mu) = \int_\pi\int_z v[\pi - s(\pi)]h_H(\pi,z)dzd\pi + \lambda\left[\int_\pi\int_z u[s(\pi)]h_H(\pi,z)dzd\pi - c(H) - \bar{u}\right] + \mu\left[\int_\pi\int_z u[s(\pi)]h_H(\pi,z)dzd\pi - c(H) - \int_\pi\int_z u[s(\pi)]h_L(\pi,z)dzd\pi - c(L)\right]$$

一阶条件为

$$\frac{\partial L}{\partial s} = \iint\left[-v' + \lambda u' + \mu u' - u'\frac{h_L(\pi,z)}{h_H(\pi,z)}\right]h_H(\pi,z)dxd\pi = 0$$

同前面的假定相同，当我们假定委托人不知道 $h_H(\pi,z)$ 时，就得到充分必要条件：

$$-v' + \lambda u' + \mu u' - u'\frac{h_L(\pi,z)}{h_H(\pi,z)} = 0$$

或

$$\frac{v'[\pi - s(\pi,z)]}{u'[s(\pi,z)]} = \lambda + \mu\left[1 - \frac{h_L(\pi,z)}{h_H(\pi,z)}\right] \qquad (8.19)$$

如果我们假定委托人知道 $h_H(\pi,z)$，则式（8.19）就是充分条件。比较条件（8.17）和（8.19），可知下列条件成立时，则新的观测量 z 是没有信息量的。

$$\frac{h_L(\pi,z)}{h_H(\pi,z)} = \frac{f_L(\pi)}{f_H(\pi)} = P(\pi) \qquad (8.20)$$

Holmstrom（1979）曾证明：当且仅当条件（8.20）不成立时，$s(\pi,z)$ 才帕累托优于 $s(\pi)$。即只有当 z 影响到似然率 $\frac{h_L}{h_H}$ 时，z 才应该进入合约。[12]

当条件（8.20）成立时，称 π 是相对于 a（和 θ）的有关 $x = (\pi,z)$ 的“充足统计量”（sufficient statistic），所有 z 能提供的有关 a（和 θ）的信息都已包含在 π，z 不提供任何额外的信息。故将 z 写进合约是没有意义的。

条件（8.20）可改写为

$$h_L(\pi,z) = h_H(\pi,z)P(\pi)$$

$$h_H(\pi,z) = h_H(\pi,z)\cdot 1$$

记 $A(\pi,z)=h_H(\pi,z)$， $B_L(\pi)=P(\pi)$， $B_H(\pi)=1$

$$h_i(\pi,z)=A(\pi,z)B_i(\pi)，\quad i=L,H \tag{8.21}$$

显然，当且仅当条件（8.21）对任何 z 成立时， π 就是充足统计量。

当条件（8.20）不成立时，将 z 写入合同 $s(\pi,z)$ 之所以有价值，是因为通过使用 z 包含的信息量，委托人可排除掉更多的外生因素对其推断的干扰，使代理人承担较小的风险（假定代理人是风险规避的），从而可节约风险成本（即使风险分摊更加接近帕累托最优水平）。节约风险或本与观测 z 的成本之间有权衡。譬如，假定分布 $h_i(\pi,z)$ 满足一阶随机占优条件和单调似然率要求，则给定代理人选 $a=H$，较低的 π 和较低的 z 同时出现的可能性显然小于较低的 π 单独出现的可能性（ $\Pr ob(\pi\le\pi_0,z\le z_0)=\Pr ob(\pi\le\pi_0\,|\,z\le z_0)\quad \Pr ob(z\le z_0)\le\Pr ob(\pi\le\pi_0)$ ）；同样，给定代理人选 $a=L$，较高 π 和较高的 z 同时出现的可能性显然小于较高的 π 单独出现的可能性。

因此，为了诱使代理人选 $a=H$，当合约只依赖于 π 时，代理人应承担的风险必大于当合约同时依赖于 π 和 z 时代理人应承担的风险（直观上，在合约 $s(\pi,z)$ 下，代理人被错误地惩罚和被错误地奖励的可能性都变得较小）。

因为代理人是风险规避的，这种由代理人承担较低风险带来的风险成本的节约就是将 z 写入合约的净价值（假定观测 z 是无成本的）。

另一方面，上述充足统计量结果对最优激励合同的设计有着重要的意义。首先，对代理人实施监督是有意义的，因为监督可以提供更多的有关代理人行动选择的信息，从而可减少代理人的风险成本。但是，监督本身也存在成本，若监督成本太高，则监督也可能是没有意义的，即使它可提供更多的信息。

更加重要的是，充足统计量结果意味着使用相对业绩比较是有意义的。比如，同一行业不同企业的经营业绩除受到每个企业经营者的行为和特有的外生因素影响外，也受到某些行业共同因素（如市场需求、技术进步等）的影响。

企业自身的利润并不是充足统计量，其它企业的利润也包含着有关该企业经理行为的有价值的信息。比如，一个企业的利润低可能是由于经理没有努力工作，也可能是由于不利的外部因素造成的。但若其他处于类似环境的企业的利润也很低，该企业利润低则更可能是不利的外部因素造成的；相反，若其它处于类似环境的企业的利润较高，该企业利润低更可能是经理不努力的结果。

通过将其它企业的利润指标引入对该企业经理的奖惩合约，则可剔除更多的外部不确定性的影响，使该经理的报酬与其个人努力的关系更为密切，从而调动其努力工作的积极性。

因此，处于类似经营环境的企业经理的报酬不应该只依赖于本企业的利润，而应部分地依赖于其它企业的利润。这就是“标尺竞争”（yardstick competition）被广泛使用的主要原因之一。

8.4.3 一般性的模型

下面，我们讨论当a是一个一维的连续变量的一般情形。在经济学上，前述a只取两种行动的模型已包含了委托—代理模型的基本结论，一般性模型并不提供多少更新的东西。但考虑一般性模型具有重要的方法论意义。此时，分布函数的一阶随机占优条件为：

$$F_a(\pi,a)=\frac{\partial F}{\partial a}<0$$

即对所有的π，若$a>a'$，$F(\pi,a)<F(\pi,a')$

激励相容条件为：

IC：对任何给定的激励合约$s(\pi)$，代理人总是选a满足：

$$\max_a \int u[s(\pi)]f(\pi,a)d\pi-c(a)$$

据 Mirrlees（1974）和 Holmstrom（1979），IC 条件导出以下一阶条件：

$$\frac{\partial IC(a)}{\partial a}=0\text{，其中 }IC(a)=\int u[s(\pi)]f(\pi,a)d\pi-c(a)$$

即$\int u[s(\pi)]f_a(\pi,a)d\pi=c'(a)$ （8.22）

此即所谓“一阶条件方法”（the first-order approach）。

由式（8.22），委托人问题变为：

$$\max_{s(\pi)}\int v[\pi-s(\pi)]f(\pi,a)d\pi$$

$$s.t\quad (IR)\quad \int u\left[s(\pi)\right]f(\pi,a)d\pi-c(a)\geq\bar{u}$$

$$(IC)\quad \int u\left[s(\pi)\right]f_a(\pi,a)d\pi=c'(a)$$

拉格朗日函数为

$$L(s,\lambda,\mu)=\int v[\pi-s(\pi)]f(\pi,a)d\pi+\lambda\left[\int u[s(\pi)]f(\pi,a)d\pi-c(a)-\bar{u}\right]$$
$$+\mu\left[\int u[s(\pi)]f_a(\pi,a)d\pi-c'(a)\right]$$

一阶条件：

$$\frac{\partial L}{\partial s}=\int\left\{-v'+\lambda u'+\mu u'\frac{f_a(\pi,a)}{f(\pi,a)}\right\}f(\pi,a)d\pi=0$$

假定委托人不知道$f(\pi,a)$，就得到充分必要条件

$$-v' + \lambda u' + \mu u' \frac{f_a(\pi,a)}{f(\pi,a)} = 0$$

当委托人知道 $f(\pi,a)$ 时，该式是充分条件。

即 $\frac{v'[\pi - s(\pi)]}{u'[s(\pi)]} = \lambda + \mu \frac{f_a(\pi,a)}{f(\pi,a)}$ （8.23）

条件（8.23）是条件（8.8）的一般化，其中 $\frac{f_a(\pi,a)}{f(\pi,a)}$ 是似然率 $\frac{f_L(\pi)}{f_H(\pi)}$ 的对应。[13]

比较式（8.23）和式（8.12）给出的帕累托最优风险分摊条件，式（8.23）意味着当委托人不能观测到代理人的努力水平 a 时，帕累托最优风险分摊是不可能的，因为 $\mu > 0$（IC 条件成立）。

为使代理人有积极性努力工作，代理人现在必须承担更大的风险。譬如

当 $\frac{f_a(\pi,a)}{f(\pi,a)} < 0$，则 $s(\pi) < s_\lambda(\pi)$

当 $\frac{f_a(\pi,a)}{f(\pi,a)} > 0$，则 $s(\pi) > s_\lambda(\pi)$

一般地，若单调似然率特征（MLRP）成立，$\frac{f_a(\pi,a)}{f(\pi,a)}$ 是 π 的单调增函数，最优激励合约 $s(\pi)$ 一定是 π 的增函数，$\frac{\partial s(\pi)}{\partial(\pi)} > 0$。

当然，在一般化模型中也存在充足统计量的问题，这里不再展开讨论。

需要指出的是：与前述 a 只取两个值的情形不同，这里当 a 为连续变量时，新的变量 z 进入合约不仅可降低风险成本，而且还可以提高努力水平（若 z 包含更多的信息）。一阶条件方法存在的一个问题，即可能带来多重解。即对于一个给定合约 $s(\pi)$，代理人的最优化条件（8.22）可能有多个解。[14]

这是因为，一阶条件即所谓："驻点必要条件"，包括了极大、极小值和所有驻点，驻点不一定就是最大值点。反过来，这就意味着最优化条件（8.22）并不能保证解就是最优的。在图 8.8 中，$s = s(\pi)$ 代表任意的激励合约，它根据委托人的偏好顺序从左到右排列。即给定 a，委托人更偏好于更靠右边的 s。满足一阶条件（8.22）的努力水平 a 由曲线 $a(s)$ 表示，它有多重解。

注意：对给定的 s，代理人选择位于虚线部分的 a，因为他会尽量少努力。

v_0 和 v_1 是委托人的两条无差异曲线（尽管 a 不直接进入效用函数 $v[\pi - s(\pi)]$，但通过分布函数 $f(\pi,a)$ 影响到期望效用）；另外，a 愈大，委托人的期望效用就会愈大，故委托人更偏好较大的 a，因而 a 与 s 之间存在替代性，无差异曲线是凸向原点的。

a

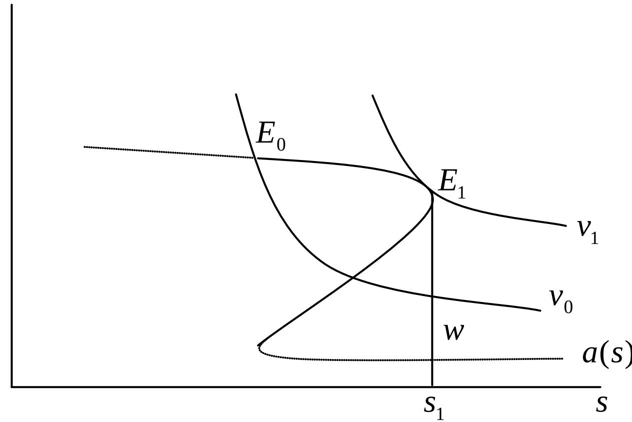


图 8.8　多重激励相容解

由一阶条件（8.23）决定的最优解为 E_1，因它在满足条件（8.22）的情况下达到最高的无差异曲线。但实际的最优解是 E_0，因它是委托人能得到的最好结果（若委托人选 s_1，代理人将选 w 而非 E_1）。因为 $a(s)$ 为单折曲线时，代理人必选其上一点时驻点必是最大值。

Mirrlees（1975）最早认识到一阶条件方法存在的上述问题。Grossman 与 Hart（1983），以及 Rogerson（1985）导出了保证一阶条件方法有效性的条件。他们证明了，若分布函数满足 MLRP 和凸性条件（CDFC，Convexity of distribution function condition），则一阶条件方法是适用的。因为在 CDFC 下，对于任何给定的 $s(\pi)$，满足一阶条件（8.22）的 $a[s(\pi)]$ 是唯一的（这里 CDFC 实际上是规模报酬随机递减的特征）。

8.6 委托—代理的 Holmstrom-Milgrom 模型

本节介绍一个重要的委托—代理模型，由 Holmstrom 与 Milgrom（1973）给出，它是一个适当简化的一维连续变量一般化模型，用参数化方法表述。该模型在文献中被大量采用。

设 a 为一维变量，$\pi = a + \theta$，θ 是均值为零、方差为 σ^2 的正态分布随机变量，是外生不确定因素。故有

$E\pi = a$，$Var(\pi) = Var(a) + Var(\theta) = \sigma^2$

其中 Var 表示方差。

假定：委托人是风险中性的，代理人是风险规避的。

现在考虑线性合约（前面的分析表明，当委托人和代理人的绝对风险规避度都为常数时，最优合约是线性的，这里假定不变的绝对风险规避度）：

$s(\pi)=\alpha+\beta\pi$，α 为代理人的固定收入，β 为代理人分享的产出份额。

当 $\beta=0$，代理人不承担任何风险

当 $\beta=1$，代理人承担全部风险

因为委托人是风险中性的，则给定 $s(\pi)=\alpha+\beta\pi$，委托人的期望效用等于

$$\begin{aligned}Ev[\pi-s(\pi)]&=v\{E[\pi-s(\pi)]\}\\&=v\{E[-\alpha+(1-\beta)\pi]\}\\&=v[-\alpha+(1-\beta)a]\end{aligned}$$

因为 v' 为常数，不妨设 $v(w)=w$，w 为委托人的收入。

则委托人的期望效用等于期望收入：

$$Ev[\pi-s(\pi)]=-\alpha+(1-\beta)a$$

现设代理人的效用函数具有不变的绝对风险规避特征。

$$\rho=-\frac{u''}{u'},\quad \rho>0$$

ρ 为代理人的绝对风险规避度，ω 为代理人的实际货币收入。

则

$$\frac{d\ln u'}{d\omega}=-\rho$$

$\ln u'=-\rho\omega+B$，B 为常数

$$u'=e^{B}e^{-\rho\omega}$$

$u=-\dfrac{e^{B}}{\rho}e^{-\rho\omega}+C$，$C$ 为常数

不妨再设 $\dfrac{e^{B}}{\rho}=1$，$C=0\quad(B=\ln\rho)$

则 $u=-e^{-\rho\omega}$

进一步简化，设 $c(a)$ 等价于货币成本[15]，具体为

$c(a)=\dfrac{ba^{2}}{2}$，这里 $b>0$ 为成本系数；

代理人的实际收入为

$$\omega=s(\pi)-c(a)=\alpha+\beta(a+\theta)-\frac{b}{2}a^{2}$$

下面，我们引入“确定性等价收入”的概念。

定义 8.1　若 $u(x)=Eu(\omega)$，ω 为随机收入，$u(x)$ 为效用函数，则称 x 为 ω 的确定性等价收入。

由此，对于代理人，我们有

$-\exp(-\rho x)=\int[-\exp(-\rho\omega)]f(\omega)d\omega$，其中 $f(\omega)$ 为 ω 的密度函数。

故 $-\exp(-\rho x)=-\int\exp\left\{\left[-\rho[\alpha+\beta(a+\theta)-\frac{b}{2}a^2\right]\right\}g(\theta)>d\theta$

$$=-\exp\left\{-\rho\left[[\alpha+\beta a-\frac{b}{2}a^2\right]\right\}\int\exp(-\rho\beta\theta)g(\theta)d\theta$$

$$=-A\exp\left\{-\rho\left[[\alpha+\beta a-\frac{b}{2}a^2\right]\right\}\int\exp\left\{-\rho\beta\theta-\frac{\theta^2}{2\sigma^2}\right\}d\theta$$

其中 $g(\theta)=A\exp\left(-\frac{\theta^2}{2\sigma^2}\right)$ 是 θ 的正态分布密度函数

故

$$-\exp(-\rho x)=-A\exp\left\{-\rho\left[[\alpha+\beta a-\frac{b}{2}a^2\right]+\frac{\rho^2\beta^2\sigma^2}{2}\right\}$$

$$\cdot\int\exp\left\{-\left[\frac{\theta}{\sqrt{2}\sigma}+\frac{\rho\beta\sigma}{2}\right]^2\right\}d\theta$$

$$=-A\exp\left\{-\rho\left[[\alpha+\beta a-\frac{b}{2}a^2-\frac{\rho\beta^2\sigma^2}{2}\right]\right\}\sqrt{2}\sigma$$

$$\cdot\int\exp\left\{-\left[\frac{\theta}{\sqrt{2}\sigma}+\frac{\rho\beta\sigma}{2}\right]^2\right\}d\left(\frac{\theta}{\sqrt{2}\sigma}+\frac{\rho\beta\sigma}{\sqrt{2}}\right)$$

$$=-\sqrt{2}A\sigma\exp\left\{-\rho\left[[\alpha+\beta a-\frac{b}{2}a^2-\frac{\rho\beta^2\sigma^2}{2}\right]\right\}\int_{-\infty}^{\infty}\exp(-y^2)dy$$

又因为 $1=\int_{-\infty}^{\infty}g(\theta)d\theta=\int_{-\infty}^{\infty}A\exp\left[-\left(\frac{\theta}{\sqrt{2}\sigma}\right)^2\right]d\theta=\sqrt{2}A\sigma\int_{-\infty}^{\infty}\exp(-y^2)dy$

故 $\int_{-\infty}^{\infty}e^{-y^2}dy=\frac{1}{\sqrt{2}A\sigma}$

即有 $-\exp(-\rho x)=-\exp\left\{-\rho\left[\alpha+\beta a-\frac{b}{2}a^2-\frac{\rho\beta^2\sigma^2}{2}\right]\right\}$

得到 $x=\alpha+\beta a-\frac{b}{2}a^2-\frac{\rho\beta^2\sigma^2}{2}$

因为 $E\omega=\alpha+\beta a-\frac{b}{2}a^2$

故 $x=E\omega-\frac{\rho\beta^2\sigma^2}{2}$

或 $E\omega-x=\frac{1}{2}\rho\beta^2\sigma^2$

$\frac{\rho\beta^2\sigma^2}{2}$ 被称为代理人的风险贴水或风险成本，即代理人在收入中放弃 $\frac{\rho\beta^2\sigma^2}{2}$ 的收

入以换取确定性收入可获同样效用，故$\frac{\rho\beta^2\sigma^2}{2}$是代理人购买保险的价格。

当$\beta=0$时，风险成本=0，此时代理人收入为确定性的。

代理人最大化期望效用函数$Eu(\omega)=-Ee^{-\rho\omega}$等价于最大化上述确定性等价收入，因为有$Eu(\omega)=u(x)$。

设$\overline{\varpi}$为代理人的保留收入水平，则当确定性等价收入$x<\overline{\varpi}$，代理人将不接受合约。故参与约束可用确定性收入表示为

$$x=\alpha+\beta a-\frac{1}{2}\rho\beta^2\sigma^2-\frac{b}{2}a^2\geq\overline{\varpi}$$

我们下面先考察一下委托人可观测到代理人努力水平a时的最优合约。此时 IC 条件不起作用，任何a都可以通过满足参与约束 IR 的一种强制性合约博弈实现，正如我们在前面所看到的那样。

此时，委托人的问题是选择(α,β)和a解下列最大化问题：

$$\max_{\alpha,\beta,a} Ev=-\alpha+(1-\beta)a$$

$$s..t\ \ (IR)\ \ \alpha+\beta a-\frac{1}{2}\rho\beta^2\sigma^2-\frac{b}{2}a^2\geq\overline{\varpi}$$

在最优情况下，IR 等式成立，故将 IR 通过固定项α代入目标函数，则上述最大化问题就变为：

$$\alpha=\overline{\varpi}-\beta a+\frac{1}{2}\rho\beta^2\sigma^2+\frac{b}{2}a^2 \qquad (IR)$$

代入Ev：

$$Ev=-\left(\overline{\varpi}-\beta a+\frac{1}{2}\rho\beta^2\sigma^2+\frac{b}{2}a^2\right)+(1-\beta)a$$

$$=a-\overline{\varpi}-\frac{1}{2}\rho\beta^2\sigma^2-\frac{b}{2}a^2$$

故$\max_{\alpha,\beta,a}\left[a-\frac{1}{2}\rho\beta^2\sigma^2-\frac{b}{2}a^2-\overline{\varpi}\right]$

因为$\overline{\varpi}$为给定的，上述表述意味着委托人实际上是在最大化总的确定性等价收入。因为委托人的确定性收入为$[-\alpha+(1-\beta)a]$（前面已有委托人风险中性假定及$v(x)=x$的假定，导出$Ev[\pi-s(\pi)]=-\alpha+(1-\beta)a$）。

总的确定性收入为$x+[-\alpha+(1-\beta)a]$

$$=\alpha+\beta a-\frac{1}{2}\rho\beta^2\sigma^2-\frac{b}{2}a^2-\alpha+(1-\beta)a$$

$$=\alpha-\frac{1}{2}\rho\beta^2\sigma^2-\frac{b}{2}a^2$$

一阶条件

$a^* = \frac{1}{b}$，$\beta^* = 0$

代入 IR 条件，得

$$\alpha^* = \overline{\omega} + \frac{b}{2}(a^*)^2 = \overline{\omega} + \frac{1}{2b}$$

这就是最优合约。

因为委托人是风险中性的，代理人是风险规避的，帕累托最优风险分摊要求代理人不承担任何风险 $(\beta^* = 0)$，而委托人支付给代理人的固定收入刚好等于代理人的保留工资 $\overline{\omega}$ 加上努力的成本 $\frac{b}{2}(a^*)^2$。最优努力水平要求努力的边际期望利润等于努力的边际成本，即 $1 = ba$。

于是 $a^* = \frac{1}{b}$。

因为委托人可观测到代理人的选择 a，只要委托人在观测到代理人选了 $a < \frac{1}{b}$ 时就支付 $\underline{\alpha} < \overline{\omega} < a^*$，代理人就一定会选 $a = \frac{1}{b}$，最优风险分摊与激励没有矛盾。

但若委托人不能观测到 a，上述帕累托最优风险分摊及努力水平都不能实现。

这是因为，给定 $\beta = 0$，代理人将选 a 最大化自己的确定性等价收入 x，因为

$$x = \alpha + \beta a - \frac{1}{2}\rho\beta^2\sigma^2 - \frac{b}{2}a^2 = \alpha - \frac{b}{2}a^2$$

一阶条件

$$a = 0$$

即若代理人收入与产出无关，代理人将选 $a = 0$ 而不是 $a = \frac{1}{b}$。

现在再看 a 是不可观测时的最优合约。

因为给定 (α, β)，代理人的激励相容约束为最大化其确定性等价收入 x：

一阶条件：$\frac{\partial x}{\partial a} = \beta - ba = 0$

$$a = \frac{\beta}{b}$$

委托人的问题是最大化其确定性收入

$$\max_{\alpha,\beta}[-\alpha + (1-\beta)a]$$

$$s..t \quad (IR) \quad \alpha + \beta a - \frac{1}{2}\rho\beta^2\sigma^2 - \frac{b}{2}a^2 \geq \overline{\omega}$$

$$(IC) \quad a = \frac{\beta}{b}$$

由 IR 条件：$\alpha = \overline{\omega} - \beta a + \frac{1}{2}\rho\beta^2\sigma^2 + \frac{b}{2}a^2$，将 IR 和 IC 条件代入目标函数，得：

$$\max_{\beta}\alpha\left[\frac{\beta}{b}-\frac{1}{2}\rho\beta^2\sigma^2-\frac{b}{2}\left(\frac{\beta}{b}\right)^2-\overline{\omega}\right]$$

一阶条件为：$\frac{1}{b}-\rho\beta\sigma^2-\frac{\beta}{b}=0$

得 $\beta=\frac{1}{1+b\rho\sigma^2}>0$ （8.24）

即代理人必须承担一定的风险。特别地，β 是 ρ、σ^2 和 b 的减函数。即代理人愈是风险规避，产出 π 的方差愈大，代理人愈是害怕努力工作，他应承担的风险就愈小。

在一个极端的情形，若代理人是风险中性的 $(\rho=0)$，最优合约就要求代理人承担完全的风险 $(\beta=1)$。

$\frac{\partial\beta}{\partial\rho}<0$ 和 $\frac{\partial\beta}{\partial\sigma^2}<0$ 的意义是很直观的。最优激励合约要在激励与保险之间求得平衡。对于给定的 β，ρ 愈大（或 σ^2 愈大），风险成本就愈高，故最优风险分摊要求 β 愈小。

但是，$\frac{\partial\beta}{\partial b}<0$ 却有点“鞭打快牛”的味道。为什么代理人愈是害怕努力工作，应该承担的风险就愈小呢？有两个方面的原因：第一，从激励看，即使没有信息不对称，b 愈大，最优的 a 就愈小（因为 $a^*=\frac{1}{b}$）；第二，从风险分摊看，b 愈大，为了诱使代理人选择同样的努力水平要求的 β 愈大（因为 $a=\frac{\beta}{b}$）。委托人宁愿以较低的努力换取风险成本的节约。既然给予代理人较大的风险也难以激励其努力增加多少，但风险成本却太高，不如为节约风险成本减少代理人承担的风险，此时代理人的努力减少不多。

当委托人不能观测代理人努力水平 a 时，存在两类在对称信息下不存在的代理成本。

一类是上面多次提到的由于帕累托最优风险分摊无法达到而出现的风险成本；另一类是由较低的努力水平导致的期望产出的净损失减去努力成本的节约，简称为激励成本。

因为委托人是风险中性的，故努力水平可观测时委托人承担全部风险意味着风险成本为零。

我们来说明为什么按上述方式定义的成本是一种成本的概念。

首先，在委托人是风险中性、代理人是风险规避者的情形，如果信息是对称的，则委托人承担全部风险，代理人不承担风险，收入是固定的。但在信息不对称的情形，代理人因激励需要而承担了一定的风险，收入变成变动的随机收入，则在同样的达到保留效用的水平上，委托人需要向代理人支付一个“保险价格”即风险贴水，这个风险贴水

就构成了由于信息不对称委托人花费的一部分成本，称为风险成本，它等于 $\frac{1}{2}\rho\beta^2\sigma^2=\frac{\rho\sigma^2}{2(1+b\rho\sigma^2)^2}>0$，记为

$$\Delta RC=\frac{\rho\sigma^2}{2(1+b\rho\sigma^2)^2}$$

于是有

定义 8.2　称风险贴水 ΔRC 为风险成本，其中 ΔRC 为

$$\Delta RC=\frac{\rho\sigma^2}{2(1+b\rho\sigma^2)^2} \tag{8.25}$$

在信息对称的情形，代理人努力水平 a 可观测，则 $a^*=\frac{1}{b}$；在信息不对称的情形，a 不可观测，委托人可诱使代理人自动选择的最优努力水平为

$$a=\frac{\beta}{b}=\frac{1}{b(1+b\rho\sigma^2)}<\frac{1}{b}=a^*$$

这是不对称信息下的最优合约诱导出的努力水平。即非对称信息下的最优努力水平严格小于对称信息下的帕累托最优努力水平。

因为 $E\pi=a$，则期望产出因信息不对称带来的净损失为

$$\begin{aligned}\Delta E\pi&=\Delta a\\&=a^*-a\\&=\frac{1}{b}-\frac{1}{b(1+b\rho\sigma^2)}\\&=\frac{\rho\sigma^2}{1+b\rho\sigma^2}>0\end{aligned}$$

但是，由于代理人在非对称信息下的努力水平较之对称信息下有所减少，委托人需要给予支付的为了保证参与约束等式成立的一部分收入也相应减少，这就扣除了委托人由于产出减少的一部分成本，这种减少额为

$$\begin{aligned}\Delta c&=c(a^*)-c(a)\\&=\frac{1}{2b}-\frac{1}{2b(1+b\rho\sigma^2)^2}\\&=\frac{2\rho\sigma^2+b(\rho\sigma^2)^2}{2(1+b\rho\sigma^2)^2}\end{aligned}$$

称 Δc 为努力成本的节约。

$\Delta E\pi-\Delta c$ 是一种净的成本，称为激励成本，于是有

定义 8.3　除 $\Delta E\pi-\Delta c$ 为激励成本，其中 $\Delta E\pi-\Delta c$ 为：

$$\Delta E\pi - \Delta c = \frac{b\rho^2\sigma^4}{2(1+b\rho\sigma^2)^2} > 0 \qquad (8.26)$$

这两种成本之和为总的代理成本。

定义 8.4　称激励成本与风险成本之和为代理人成本，记为 AC，其中有

$$AC = \Delta RC + (E\pi - \Delta c)$$
$$= \frac{\rho\sigma^2}{2(1+b\rho\sigma^2)} > 0 \qquad (8.27)$$

当代理人为风险中性者时，代理成本为零，因此时 $\rho = 0$，故 $\beta = 1$，因 $\beta = 1$ 可达到帕累托最优风险分摊和最优的激励。

$a = \frac{1}{b} = a^*$。

进一步，代理成本随代理人风险规避度 ρ 和产出方差 σ^2（刻出不确定性）的上升而上升。

下面，我们来分析可观测的其它变量如何影响最优合约。令 z 是另一个可观测的变量，为简单计，假定 z 与努力水平 σ 无关（如 z 是另一个企业的利润）,但 z 可能与外生的 θ 有关，从而 z 与 π 相关。

设 z 是正态分布，均值为零，方差为 σ_z^2 的随机变量，考虑一个线性合约（推导与前类似）。

$s(\pi, z) = \alpha + \beta(\pi + \gamma z)$

其中 β 代表激励强度，γ 表示代理人的收入与 z 的关系；

若 $\gamma = 0$，则代理人收入与 z 无关。

委托人的问题是选择最优的 α，β 和 γ。

在此合约下，代理人的确定性等价收入为：

$$\alpha + \beta a - \frac{1}{2}\rho\beta^2 \operatorname{var}(\pi + \gamma z) - \frac{ba^2}{2} \text{（推导与前类似）}$$
$$= \alpha + \beta a - \frac{1}{2}\rho\beta^2\left[\sigma^2 + \gamma^2\sigma_z^2 + 2\gamma cor(\pi, z)\right] - \frac{b}{2}a^2$$

其中 $cor(\pi, z)$ 是 π 和 z 的协方差。

对任何给定的合约 $s(\pi, z)$，代理人选 a 最大化上述确定性等价收入。

一阶条件

$$a = \frac{\beta}{b}$$

这与前面的结果相同。这是显然的，因为 z 与 a 无关，γ 不影响代理人的努力水平选择。z 的作用是为委托人提供识别 θ 的信号，不影响代理人的行为。

委托人的期望收入为

$E[\pi-\alpha-\beta(\pi+\gamma z)]=-\alpha+(1-\beta)a$

将 IR 约束和 IC 条件 $a=\dfrac{\beta}{b}$ 代入上式，得到委托人最优化问题

$$\max_{\beta,\gamma}\left\{\frac{\beta}{b}-\frac{1}{2}\rho\beta^2\left[\sigma^2+\gamma^2\sigma_z^2+2\gamma cor(\pi,z)\right]-\frac{\beta^2}{2b}-\overline{\omega}\right\}$$

得两个一阶条件

$$\frac{1}{b}-\rho\beta\left[\sigma^2+\gamma^2\sigma_z^2+2\gamma cor(\pi,z)\right]-\frac{\beta}{b}=0 \quad （8.26）$$

$$\gamma\sigma_z^2+cor(\pi,z)=0 \quad （8.27）$$

因 z 与期望收入无关（因为 $a=\dfrac{\beta}{b}$ 与 z 无关），委托人选择 γ 只是使风险成本最小化。

由式（8.26）、（8.27），得

$$\beta=\frac{1}{1+b\rho[\sigma^2-cor^2(\pi,z)/\sigma_z^2]} \quad （8.28）$$

$$\gamma=-\frac{cor(\pi,z)}{\sigma_z^2} \quad （8.29）$$

需要注意的是，由施瓦兹不等式 $\sigma^2\sigma_z^2\geq cor^2(\pi,z)$，式（8.28）分母中括号内的项为正。

故 $0<\beta<1$

现在来看式（8.28）、（8.29）的含义。

若 π 与 z 无关，$cor(\pi,z)=0$。

则 π 是充足统计量，z 不提供有关 a 的任何信息；

故 $\gamma=0$，即 z 不进入合约。

此时 $\beta=\dfrac{1}{1+b\rho\sigma^2}$，回到前面的式（8.24）。

若 π 与 z 正相关，则 $cor(\pi,z)>0$，故 $\gamma<0$，$z>0$ 可能意味着较好的外部条件（较大的 θ），在某种意义上，z 反映了 θ 的信息，任何给定的 π 可能更多地反映了代理人碰到了好运气而不是作出了高水平的努力。

类似地，$z<0$ 可能意味着较差的外部条件（较低的 θ），任何给定的 π 可能更多地反映了较高的努力水平。

故 $\gamma<0$，将所有这些可能性考虑进去：即外部因素不利时增加代理人的报酬，外部因素有利时减少代理人的报酬。

在另一方面，若 π 与 z 负相关，$cor(\pi,z)<0$。

故 $\gamma<0$，$z>0$ 更可能意味着较不利的外部环境。

$z<0$更可能意味着较有利的外部环境。

故通过在$z>0$时增加代理人的报酬，在$z<0$时减少代理人的报酬，可剔除更多的外部环境的影响。

将式（8.28）与（8.24）对比可看出：当$cor(\pi,z)\neq 0$时，通过将z写进合约，一方面可提高代理人分享的剩余分额：

$$\beta=\frac{1}{1+b\rho[\sigma^2-cor^2(\pi,z)/\sigma_z^2]}>\frac{1}{1+b\rho\sigma^2}$$

从而提高合同的激励强度（激励约束意味着$a=\frac{\beta}{b}$）。

不难看出，只要z与θ相关，将z写进合同还可以减少激励成本和代理成本。在合约$s(\pi,z)$下，风险成本为

$$\Delta RC=\frac{1}{2}\rho\,\mathrm{var}[s(\pi,z)]$$

$$=\frac{\rho[\sigma^2-cor^2(\pi,z)/\sigma_z^2]}{2\{1+b\rho[\sigma^2-cor^2(\pi,z)/\sigma_z^2]\}^2}$$

期望产出的净损失为$\Delta E\pi=\Delta a=\frac{1}{b}-\frac{\beta}{b}=\frac{\rho[\sigma^2-cor^2(\pi,z)/\sigma_z^2]}{1+b\rho[\sigma^2-cor^2(\pi,z)/\sigma_z^2]}$

努力成本的净节约为

$$\Delta c=c(a^*)-c(a)=\frac{1}{2b}-\frac{1}{2b\{1+b\rho[\sigma^2-cor^2(\pi,z)/\sigma_z^2]\}^2}$$

$$=\frac{2\rho[\sigma^2-cor^2(\pi,z)/\sigma_z^2]+b\{\rho[\sigma^2-cor^2(\pi,z)/\sigma_z^2]\}^2}{2\{1+b\rho[\sigma^2-cor^2(\pi,z)/\sigma_z^2]\}^2}$$

故总激励成本为

$$\Delta E\pi-\Delta c=\frac{\rho[\sigma^2-cor^2(\pi,z)/\sigma_z^2]}{2\{1+b\rho[\sigma^2-cor^2(\pi,z)/\sigma_z^2]\}}$$

代理成本为

$$\Delta c=\Delta RC+(\Delta E\pi-\Delta c)=\frac{\rho[\sigma^2-cor^2(\pi,z)/\sigma_z^2]}{2\{1+b\rho[\sigma^2-cor^2(\pi,z)/\sigma_z^2]\}}$$

将上述结果与合约只依赖于π相比，无论是激励成本还是总代理成本都降低了。{16}

只有当$cor(\pi,z)=0$，代理成本才与合约只依赖于π时相同。因此时π是充足统计量，z不进入合约。

一种极端的情形是：当z与θ完全相关（正或负）时

$$\sigma^2=cor(\pi,z)/\sigma_z^2$$

此时$\beta=1$，即代理人是唯一的剩余索取者，从而努力水平等于帕累托最优水平

（$a=\frac{1}{b}$）；但 $\mathrm{var}[s(\pi,z)]=0$，即代理人不承担任何风险。

即当 z 与 θ 完全相关（正或负）时，将 z（最优地）写进合约可使代理成本变成零，实现帕累托最优。

上述结果可一般化为：对于任何观测到的新的变量 x，只要 x 包含有比原有变量 π 和 z 更多的有关 a 或 θ 的信息，将 x 写进激励合约就可降低代理成本。

当然，这里的前提是观测 x 不发生另外的成本。如果观测 x 要花费成本，则只有当观测成本小于由此带来的代理成本的降低时，将 x 写进合约才是有价值的。此时，$s=\alpha+\beta(\pi+\gamma z)=\alpha+\pi+\gamma z$，而 s 为常数。

即 γz 是对 π 中随机影响的一种完全的扣除（或补偿），从而完全消除 θ 的影响。

此时，委托人可以通过观测 z“间接”观测到 θ，故支付完全由 a 决定（为常数）。事实上，委托人才是唯一的剩余索取者，而代理人的帕累托最优选择 $a=\frac{1}{b}$ 是通过委托人的奖惩得到的，并非风险激励。

注释：

[1] 有时，代理人的“行为”本质上是不可观察的，如医生是否努力为医治病人的疾病而思考。此时，我们假定观察的“成本”是无穷大。

[2] 否则，委托—代理关系就会因第三者的难以确定而不复存在。这里研究委托—代理关系中的经济学问题，只有把作出这个假设作为分析的前提。

[3] 如果农民没有其他可带来较高收入的工作选择，则地主只须给予很少的风险补偿就可诱使农民租用土地。这时，地主获得固定地租，而农民承担全部风险。这似乎正是 1949 年之前中国农民的悲惨状况。在中国数千年的历史中，农民不堪于长期受压的悲困生活而揭竿而起的故事是十分多的。

[4] 关于$G(\theta)$、$g(\theta)$被设定为“共同知识”是否恰当，我们在后面还会有讨论。

[5] 这里s是一个函数，故在数学上这是一个泛函极值问题。

[6] 按理说问题（8.7）应通过库恩—塔克定理求解，但这里我们为了减少数学上的复杂性，假定最优解在参与约束边界上取得，即在假定参与约束是等式的前提下，就可以通过拉格朗日乘数法求解问题（8.7）。

[7] 在张维迎所著《博弈论与信息经济学》（张维迎，1996）中，未对此加以说明。在那里，在$g(\theta)$是共同知识的假定下直接给出（8.10）是不严密的。

[8] 这是因为，边际替代率与无差异曲线的斜率之间只差一个负号。

[9] 假定绝对风险规避度常数是非常特别的。一般来说，若假定ρ_p和ρ_A随收入的增加而递减（即收入愈高则愈不害怕风险），则最优合约$s^*(\pi)$就是非线性的。

[10]一般地，对委托人最优的解不一定就是$a=H$，因为倘若为使$a=H$的成本—s太大，委托人宁愿选$a=L$，但为了得到$a=L$，委托人只需规定这样一种简单的合约$s(\cdot)\equiv \underline{s}(\underline{s}$充分小$)$即可，因$s\equiv \underline{s}$时，代理人必选$a=L$。

[11]国有企业经理可能担心企业赚了钱导致其位置被上级或其他竞争者所觊觎。上级可以通过将其“提升”而调往其它虚职，或者觊觎者想方设法找出经理的“阴暗面”（如贪污、生活作风不正）而将其打倒后取而代之。当然，将企业“做亏”也可能不利于企业经理，但对于国有企业来说，经理们很容易将亏损的原因归咎于许多其它的不可控因素而推卸掉自身的责任。

[12]帕累托优于（Pareto-dominate）是指：给定代理人的参与约束和激励相容约束，委托人在$s(\pi,z)$下的期望效用严格高于在$s(\pi)$下的期望效用。

[13]对式（8.23）可作出与式（8.8）类似的经济解释。

[14]就象一般的驻点有多个解，或驻点不一定就是最大值点的道理一样。

[15]我们在前面将$c(a)$设为效用等价的成本，而这里又将$c(a)$设为货币等价的，两

者不矛盾。在这里，参与约束被表述为代理人的收入不小于保留收入，它是等价于原有的代理人效用不小于保留效用的。见接下来的推理。

[16]风险成本并未降低，但一些教科书（如张维迎（1996））却如此认为。不难验证，风险成本没有降低。同时，一些教科书还认为代理人的风险也会因 z 的进入合约而下降（如张维迎（1996）），但不难验证这种说法是错误的。

第九章　委托——代理理论：基本理论的扩充

在第八章中，我们构筑了委托——代理理论的基础。这种基础性的理论研究解决了在信息不对称情形下，委托人如何设计最优的合约去激励代理人最大程度地符合委托人的利益完成行动的选择。理论的要点是：在信息不对称的情形，与信息对称的情形相比较，委托人会以损失一定的收入作为代价去维系委托——代理的关系，即使委托人设计出最优的委托——代理合约，即存在大于零的代理成本。

在委托——代理理论随后的进一步发展中，研究者引入了代理人之间存在竞争以及博弈从一次性扩展到多阶段的重复博弈等因素，这种考虑被证明可以进一步降低代理成本。除此之外，更多的假定情形导出了委托——代理关系下更加丰富多彩的结果，这些结果反映了一个经济组织内部复杂的人类行为。在本章及下一章中，我们将概要性地对委托——代理理论的这些发展作一个介绍。基于这样的理论发展，委托——代理理论的许多模型已经开始成为现代人力资源经济学中的核心内容。

委托——代理理论的扩展是在放松第八章中的多个潜在假定的基础进行的。譬如，我们在第八章中实际上对博弈做出以下几个方面的潜在假定：委托人与代理人进行一次性博弈；代理人只有一个；代理人只从事一项工作等等。我们在本章及下一章中，通过将这些限定逐步放松来获得有关结果的修正或扩充。

9.1 委托——代理的动态模型

在前一章中，委托人与代理人之间进行的是一次性静态博弈。在这种静态博弈中，如果委托人不能“经济地”观测到代理人的行动选择，那么，为了激励代理人努力工作，委托人只能通过与代理人签订合约，通过某些双方都可观测的指标来对代理人行为进行奖惩的方式去激励代理人。这种通过基于双方都可观测的指标来形成合约的方式被称为“显性激励机制”（explicit incentive mechanism）。

我们知道，一旦博弈从静态博弈改变为动态博弈，均衡行为可能出现根本的改变。譬如，静态的“囚徒困境”博弈只有唯一的均衡是不合作的，但在动态的“囚徒困境”重复博弈中，就会出现因声誉的需要而形成的合作性均衡。我们将一般性的声誉模型的思想应用于委托——代理关系，就可以获得相应的结论，即委托人与代理人之间尽管不签订显性激励合约，也可能由于代理人需要声誉而努力工作，从而降低道德风险或代理

成本。

当博弈变成多阶段的动态博弈时，委托——代理关系的动态博弈分析告诉了我们以下的结果：

如果委托——代理关系是重复的无限次博弈，则当双方都有足够的耐心时（即贴现因子足够大），帕累托一阶最优风险分摊与激励都可同时实现。这种结果出现的机理是：在长期，大数定理会消去外生不确定性，委托人可以从众多的可观测变量中准确推断出代理人的真实努力水平，代理人在长期不可能隐瞒自己的人偷懒行为。[1]Radner(1981)和 Rubbinstein(1979)的工作证明了这一直观结果。

即使不考虑委托人与代理人之间的合约的可执行性，出于声誉的考虑，代理人也许仍然会努力工作。在这些模型中，“时间”作为一种“可观测变量”起到了监督代理人（也包括委托人）的作用。按照这种思路，Lazear(1979)研究了这样一种模型，即在长期的雇佣关系中，“工龄工资”机制可以遏制员工的偷懒行为。即在工作的前期阶段支付的工资低于员工边际生产率，这种差额起到了“保证金”的作用，一旦员工被发现偷懒，企业就立即开除他。潜在的保证金损失会激励员工努力工作，因为它增大了偷懒的成本。当然，在员工的后期工作阶段，由于衰老等因素，员工的边际生产率可能会下降，又由于工作后期较高的工资率，员工的边际生产率会低于工资率，此时已没有人愿退休，故企业会实行强制性退休制度。这个模型解释了为何在工作早期工资较低以及强制退休制度，它并非通过显性激励机制而是通过一种“保证金”制度来激励员工。所以，在“时间”因素引入之后，解决代理问题可能有许多可以想象得到的做法，完全不必仅限于显性激励合约。关于将委托——代理扩充到重复性动态博弈的工作，这里不打算作具体的介绍，读者可直接参观 Radner 和 Rubbinstein 的论文。在这里，只给出另外的两种动态博弈模型，即代理人声誉模型和棘轮效应模型。

9.1.1 代理人声誉模型

如果在代理人的就业上存在竞争，譬如，当经理是企业在竞争性的经理人市场中聘任而来的时候，经理在工作中不仅关心企业给予他的货币收入（来自显性激励），而且还可能关心其它方面的一些获得。我们现在来考虑一下其它方面的获得可能包括哪些内容。经理可能并不认为他会在一个特定的企业干一辈子，他在今后可能在就职方面还会有更多的选择。这种选择可能源于经理在职业生涯上的进取心，也可能是由于企业的未

来不可预期的变动所导致的经理就业变动，如因企业破产导致的经理再择业或企业有更好的经理人选而解聘经理。由于这方面的考虑，经理人在经理人市场上形成一种“高素质”或“高能力”的声誉可能对于经理人的未来是十分重要的。即使不考虑再择业的情形，经理人在企业还总是希望能在薪金上有所增加或职位提升。这也有赖于董事会对他的高能力的确认。特别地，在职位聘任上，无论是再择业还是原有企业中的职位提升，企业关心的一个要素是经理人的能力水平。任何产出都是经理人努力水平，经理人能力水平以及随机因素综合作用的结果。企业只能观测到产出水平，经理人的努力水平及能力水平是不可能观测的。经理人为了向企业显示自己的能力水平，就可能通过努力工作使产出指标更多地向企业表明他是一个高能力的经理人。在分离均衡，只有高能力的经理人才会努力工作，低能力者退出经理人市场。在混同均衡、高能力和低能力经理人都会努力工作。这样，经理人在工作中除了关心即期的货币收入之外，还可能关注于其个人能力的市场显示，这就是声誉。也就是说，经理在工作中还在“生产”其个人声誉。如果工作本身还会给代理人带来快乐，代理人也会努力工作的。这些都是可能添加的要素，从而使得在没有显性激励的情况下代理人也会努力工作。我们称这类情形为“隐性激励”（implicit incentive mechanism）。

Fama(1980)提出了代理人声誉模型的思想，他的观点与 Radner 及 Rubbinstein 的观点还有所区别。Fama 认为，现实中的激励问题并没有委托——代理理论中所提到的那么严重，“时间”可以解决问题。他强调代理人市场对代理人行为的约束。譬如，在竞争性经理人市场上，经理人的市场价值（收入）决定于其过去的经营业绩，在长期，经理人会对自己的行为负完全的责任。如果我们将委托——代理关系扩充到长期的多阶段重复博弈，则代理人必须关注未来收入。这样，即使没有显性激励，代理人也会努力工作，因为这样可以改进自己的声誉，并由此提高自己的未来收入。

(1)数学模型

Fama 的思想后来由 Holmstrom 完成了模型化工作（Holmstrom，1982）。下面，我们介绍代理人关注于声誉所带来的隐性激励模型，它是对 Hlomstrom 模型的一个简化的情形。

假设博弈有两个阶段，用 $t=1,2$ 表示，单阶段生产函数为：

$$\pi_t = a_t + \theta + u_t \qquad t=1,2$$

其中，π_t 为产出，a_t 是经理在阶段 t 的努力水平；θ 是经理人的能力水平（假定它

不随时间而变），u_t是外生的随机变量。

假定a_t是经理人的私人信息，而π_t是可观测变量，假设θ对于经理人市场来说也是不可观测的，但对于经理人市场是一种贝叶斯意义上的随机变量。假定θ和u_t是独立的正态分布随机变量，均值都为零，$E\theta = Eu_t = 0$，方差分别为σ_θ^2和σ_u^2；还假设u_1与u_2是独立的，即有$\text{cov}(u_1, u_2) = 0$。设$c(a)$是经理人的单阶段“负效用”或努力的成本函数，且$c'(0) = 0$，$c''(a) > 0$。

假定经理是风险中性者，且贴现金为零（贴现因子为 1）。故经理人的效用函数为

$$U = w - c(a_1) - c(a_2)$$

其中w为经理人的总收入。

根据假设，$w = w_1 + w_2$，其中w_1和w_2分别为经理在阶段 1 和 2 的工资收入。

我们先来看在单阶段博弈中，如果博弈是一次性的，其最优努力水平为何。根据第八章中的式（8.16），我们有

$$Eu'\left[\frac{\partial s^*}{\partial \pi}\frac{\partial \pi}{\partial a} - \frac{1}{u'}\frac{\partial c}{\partial a}\right] = 0$$

其中s^*是一次性委托——代理博弈中的最优合约。

在代理人风险中性假设下，由 8.4.2 有$s^* = \pi - y$。其中π为产出，y_0为委托人的固定收入，全部风险由风险中性的代理人承担。在这里的二阶段动态博弈中，如果第一阶段委托人与代理人签订显性激励合约，则最优合约就为$s^* = \pi_1 - y$。据我们在这里假定的生产函数，有

$\dfrac{\partial \pi}{\partial a} = 1$，于是有

$$E\left[1 - \frac{\partial c}{\partial a}\right] = 0$$

即　　$c'(a_1) = 1$　　（9.1）

式（9.1）决定的a_1就是第一阶段中的最优努力程度，记其为a_1^t。

下面，我们假设在这里不存在显性激励合约。比如，当经理人与企业所有者之间在可观测的产出π_t的具体计划结果上不一致时，显性合约就无法签订。[2]

在不存在显性激励机制的一次性委托——代理博弈中，经理人必定不会有任何的努力，故$a_t = 0$。这是因为，即使经理获得与产出无关的固定的收入，由于努力水平的不

观测性，经理也不会付出一丁点工作努力的。在二阶段的动态博弈中，第二阶段经理没有必要再努力工作，因为声誉已经由第一阶段的努力工作“生产”出来了，故$a_2 = 0$。但在第一阶段，经理为了“生产”在第二阶段的声誉而会努力工作，因为经理在第二阶段的工资收入w_2与经理市场或股东对经理能力的评价有关。[3]

由于产出是经理的个人努力水平，能力高低和随机因素共同作用的结果，因而产出也是随机变量。根据竞争性经理市场的边际生产率定价规则，在第一阶段，经理人的工资率等于第一阶段产出的期望值，第二阶段经理人的工资率也等于第二阶段的产出期望值，但由于两个阶段不是完全相互独立的，据我们的假设，经理的能力水平在两个阶段是相同的。因此，在第二阶段，经理市场应该根据第一阶段产出所提供的信息去捕获有关经理人能力水平的信息。这样，第二阶段经理人的工资率应等于给定第一阶段产出的情况下，第二阶段产出水平的期望值。于是，我们有

$$\begin{aligned} w_1 &= E(\pi_1) = E(a_1) = \bar{a}_1 \\ w_2 &= E(\pi_2/\pi_1) \end{aligned} \tag{9.2}$$

进一步有：

$$w_2 = E(\pi_2/\pi_1) = E(a_2/\pi_1) + E(\theta/\pi_1) + E(u_2/\pi_1) = E(\theta/\pi_1) \tag{9.3}$$

这是因为$a_2 = 0$，u_2与u_1，a及θ无关，故a_2与π_1无关。

在均衡时，经理市场知道a_1与θ的关系，并根据θ的分布能计算是$\bar{a}_1$。于是，在均衡状态，一旦观测到π_1，经理人市场就可计算出$\pi_1 - \bar{a}_1 = \theta + u$。但是，经理市场不能将$\theta$与$u_1$区分开来，经理市场的问题是通过观测到的$\pi_1$来推断$\theta$。

这是一个经济学中经常用到的统计推断问题。我们在本书的附录 1 中，给出了一个定理，根据这个定理，统计推断的公式为

$$\begin{aligned} E(\theta/\pi_1) &= (1-\tau)E(\theta) + \tau(\pi_1 - \bar{a}_1) \\ &= \tau(\pi_1 - \bar{a}_1) \end{aligned} \tag{9.4}$$

这是因为$E(\theta) = 0$。

其中

$$\begin{aligned} \tau &= \frac{Var(\theta)}{Var(\theta) + Var(u_1)} \\ &= \frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_u^2} \end{aligned} \tag{9.5}$$

Var 表示方差。

式（9.4）的含义是：给定π_1下经理市场预期的θ的期望值等于先验的期望值$E\theta$与事后观测值$(\pi_1-\bar{a}_1)$的加权平均值，经理市场在根据观测到的信息修正对经理能力水平的判断。

由于$\tau>0$，故$w_2=E(\theta/\pi_1)=\tau(\pi_1-\bar{a}_1)$，即第一阶段的产出$\pi_1$愈高，第二阶段的工资率$w_2$就愈高。将$w_1$和$w_2$代入经理的效用函数。

$$\begin{aligned}U&=\bar{a}_1-c(a_1)+\tau(\pi_1-\bar{a}_1)-c(a_2)\\&=\bar{a}_1-c(a_1)+\tau(a_1+\theta+u_1-\bar{a}_1)-c(a_2)\end{aligned} \quad （9.6）$$

显然，最大化U的一阶条件为：

$\frac{\partial U}{\partial a_1}=0$，$a_2=0$

即[4]

$$c'(a_1)=\tau$$

$$a_2=0$$

因$\tau>0$，$c'(0)=0$，$c''(a_1)>0$，故$a_1>0$ （9.7）

又由于$\tau\le 1$，故由$c'(a_1)$的严格递增性，有$0<a_1<a_1^+$。

这就是说，出于声誉的考虑，经理在第一阶段会努力工作，这在一次性博弈中是不可能的；但是，第一阶段的努力水平小于帕累托最优水平。

据式（9.5），τ是θ的方差与π_1的方差的比率。σ_θ^2愈大，τ就愈大。τ反映了π_1所包含的有关θ的信息：τ愈大，π_1包含的信息量就愈多。特别地，倘若没有事前不确定性，即$\sigma_\theta^2=0$，则$\tau=0$；当事前不确定性很大，即$\sigma_\theta^2\to\infty$，或者没有外生的不确定性，即$\sigma_u^2=0$，则$\tau=1$。所以，τ愈大，π_1传递的信息就愈多，经理人就愈愿意在第一阶段努力工作，从而在第二阶段形成愈强的声誉，故a_1是τ的增函数。[5]

注意，a由$c'(a_1)=\tau$，a_1与θ无关，故$\bar{a}_1=a_1$。

这个模型可以推广到多阶段博弈的情形。一般地，当博弈有T个阶段时，除最后一阶段a_T为零外，所有$T-1$阶段之前的各个阶段中努力水平a_t都为正，但$a_1>a_2>\cdots>a_{T-1}>a_T$，即努力随年龄的增加而下降。因为愈是接近于退休的年龄，努力的声誉效应就愈小。因为阶段1的努力a_1影响后面$(T-1)$个阶段的工资，但$(T-1)$的努力a_{T-1}只影响w_T。这解释了在公司里面为何年轻经理的工作更加卖力的原因。

进一步，当θ_t服从随机行走[6]（random walk），当$T \to \infty$时，可证明稳态一阶条件为：

$$c'(a) = \frac{\delta\tau}{1-(1-\tau)\delta}$$

其中δ为贴现因子。当$\delta = 1$时，$c'(a) = 1$，即当经理对未来不贴现时（未来与现在一样重要），帕累托最优努力水平就可实现。

(2)相对业绩比较效应

在第八章中，我们发现使用“相对业绩比较”，即当代理人收入不仅依赖于自己的业绩，而且还依赖于其他代理人的业绩时，可以增大激励效应，减少代理成本。我们自然会提出这样一个问题：类似的结论在动态博弈中是否仍成立？Meyer 与 Vickers（1994）的研究发现，结论是不确定的。

设经理i和j有相同的偏好与技术，u_{tj}是经理j在阶段t中面临的外部不确定因素，$\eta = \text{cov}(\theta_i, \theta_j)$是他们之间的能力水平的相关系数，$\rho = \text{cov}(u_{ti}, u_{tj})$。根据充足统计量的研究，只有当$\eta$和$\rho$中至少有一个不为零时，将经理$j$的产出$\pi_{tj}$作为决定经理$i$的报酬$w_{ti}$的一个可预测变量才是有意义的。因为若$\eta = \rho = 0$，$\pi_{tj}$实际上不提供关于经理$i$的任何信息。

在没有显性合约的情形，有$a_{2i} = a_{2j} = 0$，$w_{2i} = E(\theta_i / \pi_{1i}, \pi_{1j})$，即经理市场根据两个经理在过去的产出来决定对经理的能力水平的预期。根据本书附录 2 上的定理，我们有

$$\begin{aligned} w_{2i} &= E(\theta_i / \pi_{1i}, \pi_{1j}) \\ &= \frac{\tau}{1-k^2}(1-\eta k)(\pi_{1i} - \bar{a}_{1i}) + (\eta - k)(\pi_{1j} - \bar{a}_{1j}) \end{aligned} \quad (9.8)$$

其中$k = (1-\tau)\rho + \tau\eta$，它是$(\pi_{1i} - \bar{a}_{1i})$与$(\pi_{1j} - \bar{a}_{1j})$之间的“交叉相关系数”，而$\tau$由式（9.5）定义。代入经理$i$的效用函数

$$\begin{aligned} U &= w_{1i} + w_{2i} - c(a_{1i}) - c(a_{2i}) \\ &= \bar{a}_{1i} + \frac{\tau}{1-k^2}(1-\eta k)(a_{1i} + \theta + u_{1i} - \bar{a}_{1i}) \\ &\quad + (\eta - k)(\pi_{1j} - \bar{a}_{1j}) - c(a_{1i}) - c(a_{2i}) \end{aligned}$$

这里，仍用θ表示经理i的能力类型，u_{1i}是经理i在第一阶段的产出函数随机变量。

最大化U的一阶条件为：

$$c'(a_{1i})=\frac{\tau}{1-k^2}(1-\eta k) \qquad (9.9)$$

只有当$\frac{\tau}{1-k^2}(1-\eta k)>\tau$时，相对业绩比较才带来经理$i$努力的帕累托改进。这等价于$(1-\tau)k(\rho-\eta)>0$，由式（9.5）知一般有$\tau<1$，故当且仅当$k(\rho-\eta)>0$时，相对业绩比较才带来经理$i$的努力的增加。假定$\rho\geq 0$，$\eta\geq 0$，且$k=(1-\tau)\rho+\tau\eta>0$。则上述条件为$\rho>\eta$。若$\eta=0$且$\rho>0$，即能力不相关但外生因素相关，对相对业绩比较可以改进激励，因为从π_{2j}的观测中可以增强对经理i能力的判断，从而减少将高产出归于外生因素或低产出归于能力低的错误，这类似于静态博弈中的情形。若$\eta>0$且$\rho=0$，则会弱化激励，因为当观测到经理j的高产出时，相对业绩比较可能将经理 1 的高产出部分归于外生因素，但实际上经理j和i的产出随机因素不相关，经理i的能力应与高产出的经理j的高能力相关，因而经理i可能的高能力却会因此而被贬低。当经理i意识到这种可能性时，就不会努力工作了。此时，观测π_{1j}实际上弱化了π_{1i}的声誉效应。

9.1.2 棘轮效应模型

当代理人的能力愈高时，委托人支付给代理人的报酬就愈高。因此，代理人在过去的工作历史中，就有积极性通过努力工作取得好的成绩来向委托人传递他的高能力信息，或者通过好的成绩影响未来委托人对他的能力的评价，从而提高未来收入。这就是代理人声誉模型的内含。但是，如果我们将企业自身的“固有生产能力”也纳入影响企业产出的因素考虑之中，结论将会有所不同。企业的高产出不仅与经理的能力和努力水平有关，而且还与企业自身的品质有关，如资金规模，技术水平以及企业所处的行业发展情况等。因此，有可能高产出并不完全是来自经理的高能力或高努力水平，而且还来源于企业的高品质，我们称企业的品质为企业的“固有生产能力”。

由于企业的所有权属于委托人，高产出中来自于企业“固有生产能力”的贡献自然属于委托人，只有来自经理高能力和努力贡献的产出才归经理所有。但是，在信息不对称的假定下，企业的固有生产能力也与经理的个人能力以及努力水平一样，是不为委托人所知的。委托人只有通过过去的产出信息通过后验性的概率修正来估计出企业的固有生产能力。可以料想到，当企业在过去的产出愈高时，委托人在未来对企业的固有生产

能力的估计就愈高。这样一来，经理就没有积极性在过去努力工作了，因为努力工作的结果可能是通过高产出提高委托人在未来对企业固有生产能力的估计，而同时低估经理的个人能力，从而减少经理的未来收入水平。这是与前一小节中的声誉模型所得出的结论完全相反的结论。

在一般的情况下，我们面临一个如何准确评估代理人工作能力及努力水平的方法的问题。由于代理人的产出是其能力、努力水平与外部的环境（如企业固有生产能力）等因素共同作用的结果。所以，仅凭产出不能完全准确地测度代理人的能力和努力水平。从激励角度看，委托人支付给予代理人的报酬只有与代理人的能力和努力水平相关时，才会产生激励效应（或甄别功能）。但是，委托人也只能从可观察到的产出中获取信息，才可能对代理人的能力或外部环境对产出的贡献进行估计。当委托人提高对外部环境对产出的贡献的估计时，代理人的未来收入就会被压低。所以，一个实际需要解决的问题是如何设计对代理人业绩进行评估的“标准”，当代理人的业绩高于这个标准时，就可以认为代理人的能力或努力水平是“较高”的。建立业绩标准的一种方法是所谓“时间——动作研究”。另一种方法是用其他人的业绩作“相对业绩比较”。但在许多场合，这两种方法都是不适用的。譬如，当代理人是销售经理时，由于不同市场环境的情况差异很大，很难规定出一个业绩标准的销售量。对于企业总经理来说，由于管理上的个性特点和企业面临的环境的特殊性，这两种方法也难适用。此种，可替代的一种方法就是将代理人过去的业绩作为标准。譬如，代理人过去的业绩较好，今后就以这种较好的业绩作为标准，倘若代理人在今后达不到这种标准，就会受到惩罚。但是，代理人过去的业绩与代理人过去的努力有关，还不完全仅与企业固有生产能力和个人能力相关。这种以过去业绩作为标准的方法，是将代理人过去努力的成就作为未来产出的起点，就削弱了代理人在过去努力的积极性。这就好象驾车人总是将皮鞭落在跑得快的的身强力壮的牛身上，称为“鞭打快牛”或“棘轮效应”（ratchet effects）。魏茨曼（1980）在对前苏联计划经济制度的研究时，最早提出了“棘轮效应”的概念。在计划经济中，企业的年度生产指标被上级根据上一年的实际生产不断向上调整，导致努力生产反而受到惩罚，而聪明的企业经理会通过隐瞒生产能力的方法来对付上级。在西方的市场经济中，这种现象也类似存在，如政府对自然垄断性企业的价格管制中，企业的生产成本愈低，价格就愈低，这是对降低成本的惩罚。

(1)数学模型

“棘轮效应”会弱化激励机制，下面对此进行分析。

假定博弈仍是二阶段的动态博弈，生产函数与前一小节基本相同：

$\pi_t = a_t + \theta + u_t$，　$t = 1,2$

不同的一点是，这里我们将θ理解为企业的“固有生产能力”（intrinsic productivity），而非经理能力水平。设θ是服从正态分布且均值为$E(\theta) = \bar{\theta} > 0$，方差为$\sigma_\theta^2$的贝叶斯随机变量，而其它的变量及随机性质与前小节相同。委托人在每一阶段结束时可观测到产出π_t，但不能观测到θ和u_t。委托人根据观测到的π_t对θ的判断进行修正。与前面的式（9.4）一样，我们有

$$E(\theta/\pi_1) = (1-\tau)\bar{\theta} + t(\pi_1 - \bar{a}_1)$$

假定经理是风险中性者，效用函数为

$$U = w_1 - c(a_1) + w_2 - c(a_2)$$

这里与前小节不同的是，我们假定委托人与代理人在每一阶段都签订显性合约，但委托人不与代理人签订跨越二个阶段的长期合约，即第二阶段合约在委托人观测到π_{1i}后才制定。此时，最优激励机制设计要求经理承担全部风险，委托人在每一阶段只收取一个固定的“上缴额”记为α_t，即有：

$$w_t = \pi_t - \alpha_t \qquad (9.10)$$

在第一阶段，企业固有生产能力的期望值为

$$E(\theta) = \bar{\theta}$$

它应是企业“上缴”给委托人的数额，因为企业的所有权归委托人，即有$\alpha_1 = E(\theta) = \bar{\theta}$。同样有：

$$\alpha_2 = E(\theta/\pi_1) = (1-\tau)\bar{\theta} + \tau(\pi_1 - \bar{a}_1)$$

代入式（9.10），再代入代理人效用函数：

$$\begin{aligned} U &= \pi_1 - \alpha_1 - c(a_1) + \pi_2 - \alpha_2 - c(a_2) \\ &= a_1 + \theta + u_1 - \bar{\theta} - c(a_1) + a_2 + \theta + u_2 - (1-\tau)\bar{\theta} - \tau(a_1 + \theta + u_1 - \bar{a}_1) - c(a_2) \\ &= [(1-\tau)a_1 + (1-\tau)\theta + (1-\tau)u_1 - c(a_1)] + a_2 + \theta + u_2 - c(a_2) + (2-\tau)\bar{\theta} + \tau\bar{a}_1 \end{aligned}$$

一阶条件[7]

$\dfrac{\partial U}{\partial a_1} = 0$，$\dfrac{\partial U}{\partial a_2} = 0$，得到：

$$c'(a_1) = 1 - \tau < 1$$

$c'(a_2)=1$ （9.11）

这说明：经理在阶段 2 的努力水平是帕累托最优的但在阶段 1 的努力水平小于帕累托最优水平。其中的道理是：在阶段 2，代理人是唯一的剩余索取者，因为委托人只收取一个固定的数额，所以，努力的边际收益为 1，它等于边际成本 $c'(a_2)$。但在阶段 1，代理人在增加边际上一单位努力时的收益为 1 减去委托人在未来对企业固有生产能力期望值的提高 τ 单位从而使“上缴额” α_2 提高的 τ 单位，即净边际收益为 $1-\tau<1$，它等于边际成本 $c'(a_1)$。

这就是“棘轮效应”的发生机制，其作用与声誉模型相反，在声誉模型中，代理人的能力水平 θ 的不确定性愈大，则 τ 愈大（见式(9.5)），激励效应就愈大。但在棘轮效应模型中，企业固定生产能力 θ 的不确定性愈大，τ 愈大，激励效应却愈小（见式(9.11)）。

(2)模型的博弈论含义

棘轮效应产生的原因是因为显性合约是短期的，即委托人（或代理人）不能承诺长期合约。显然，委托人在阶段 1 开始时的期望收入为 $\alpha_1+\alpha_2=\bar{\theta}+\bar{\theta}=2\bar{\theta}$。如果合约是长期的，在博弈开始时就规定 $\alpha_1=\alpha_2=\bar{\theta}$，且能保证合约能得到遵守，则代理人效用为

$$\begin{aligned}U&=w_1-c(a_1)+w_2-c(a_2)\\&=\pi_1-\alpha_1-c(a_1)+\pi_2-\alpha_2-c(a_2)\\&=a_1+\theta+u_1-\alpha-c(a_1)+a_2+\theta+u_2-\alpha_2-c(a_2)\end{aligned}$$

一阶条件：

$$\frac{\partial U}{\partial a_1}=1-c'(a_1)=0$$

$$\frac{\partial U}{\partial a_2}=1-c'(a_2)=0$$

即得到二个阶段的帕累托最优努力水平所满足的条件：$c'(a_1)=c'(a_2)=1$。

再设具体地有 $c(a_t)=a_t^2/2$，则帕累托最优努力水平 $a_1^*=a_2^*=1$。

代理人的期望效用水平为：

$$\begin{aligned}EU&=a_1^*+\bar{\theta}-\bar{\theta}-c(a_1^*)+a_2^*+\bar{\theta}-\bar{\theta}-c(a_2^*)\\&=2-2\times\frac{1}{2}=1\end{aligned}$$

如果委托人不能承诺长期合约，则 $a_1^*=1-\tau$，$a_2^*=1$，代理人的效用水平为：

$$EU = 1-\tau-(1-\tau)^2/2+1+\bar{\theta}-E[E(\theta/\pi_1)]-1/2$$
$$= 2-\tau-(1-\tau)^2/2+\bar{\theta}-(1-\tau)\bar{\theta}-\tau\bar{\theta}-\frac{1}{2}$$
$$= 1-\tau^2/2<1$$

此时，委托人期望收入为

$$\bar{\theta}+E[E(\theta/\pi_1)]$$
$$=\bar{\theta}+(1-\tau)\bar{\theta}+\tau\theta=2\bar{\theta}$$

与长期合约时相同。

所以，$1-(\tau^2/2)>0$ 是短期合约与长期合约相比时的净福利损失。

尽管长期合约是一种帕累托最优的合约，但却是不能实现或难以得到遵守的，因为它不满足我们在第五章中提到的“动态一致性”，即不是精炼均衡。

在长期合约中，如果是长期合约，委托人在阶段 2 的固定收入为 $\alpha_2=\bar{\theta}$。但其到了阶段 2，给定 π_1，预期的企业固有生产率为 $E(\theta/\pi_1)=(1-\tau)\bar{\theta}+\tau(\pi_1-a_1)$。因此，若有 $(1-\tau)\bar{\theta}+\tau(\pi_1-\bar{a}_1)>\bar{\theta}$，即 $(\pi_1-\bar{a}_1)>\bar{\theta}$，则通过重新签约，委托人可增加自己的收入，则此时委托人就有积极性如此做。反之，若 $(\pi_1-\bar{a}_1)<\bar{\theta}$，则代理人也有积极性要求重新签约，这样可减少上缴的额度 α_2。所以，每一方都不会相信对方会遵守合约。a_1 每增加一个单位，期望的上缴额 α_2 就会增加 τ 个单位，即当 $(\pi_1-\bar{a}_1)>\bar{\theta}$ 时，多上缴 τ 单位，当 $(\pi_1-\bar{a}_1)<\bar{\theta}$ 时，就少缴 τ 单位。

在这种情况下，代理人不会有完全的积极性。这是因为，如果代理人的参与约束不起作用，则只有委托人有权修改合约，但由于收入的刚性，代理人收入不能比原合约规定的低。

那么，只有当 $(\pi_1-\bar{a}_1)>\bar{\theta}$ 时，合约才会被修改。此叶，增加努力会提高合约被修改的可能性。另一方面，如果只有代理人有权提出修改合约，并且只有当 $(\pi_1-\bar{a}_1)<\bar{\theta}$ 时才有理由提出，那么，增加努力会降低合约被修改的可能性。在这两种情形，代理人都不可能有完全的激励。只有当任何一方都无权修改合约时，代理人才会有完全的激励。

在现实生活中，为了得到长期合约带来的好处，双方可能都会可能尽力建立和维护“说话算数”的声誉，此时棘轮效应就不重要了。最后一个问题是，“相对业绩比较”的引入会怎样影响棘轮效应的呢？Meyer 与 Vickers（1994）曾证明，当引入相对业绩比较后，代理人的最优努力水平 a_{1i} 满足如下一阶条件：

$$c'(a_{1i}) = 1 - \tau\left(\frac{1-\eta k}{1-k^2}\right) = 1 - \rho < 1 \qquad (9.12)$$

其中 $\rho = \tau\left(\frac{1-\eta k}{1-k^2}\right)$

将式（9.12）与式（9.11）比较，我们发现，在棘轮效应模型中，与声誉模型不同之处在于：当且仅当 $k(\rho-\eta) < 0$ 时，引入相对业绩比较才强化激励机制。若假设有 $\rho > 0$，$\eta > 0$，则该条件为 $\rho < \eta$。在直观意义上，若 $\rho < \eta$（企业固有生产率之间的相关度大于外生的随机变量之间的相关度），则引入相对业绩比较将会降低第一阶段的业绩在推断内在生产率上的权数，从而弱化棘轮效应。倘若 $\rho > \eta$，则会出现相反的情形。

9.1.3 动态模型的小结

在声誉模型中，委托人根据代理人过去的业绩推断经理的能力，将会强化激励机制，但在棘轮效应模型中，委托人根据代理人过去的业绩推断企业的固有生产能力将会弱化激励机制。在引入相对业绩比较后，在两类模型中形成的对激励的影响正好相反。

两类模型中激励效应的不同源于过去业绩传递的信息有不同的所有权。声誉模型中，过去业绩传递经理能力的信息，高能力的所有权属经理。经营业绩愈好，市场就判断其经营能力就愈高，因而经理的报酬就愈高，故经理努力水平就愈大。当引入相对业绩比较后，就弱化了经理自己的业绩在评价能力上的作用（当 $\rho < \eta$ 时），因而就会弱化激励机制。反之，在棘轮效应模型中，过去的业绩所传递的信息是关于企业固有生产能力的，而它又属于委托人所有。经营的业绩愈好，委托人就认为企业的固有生产能力愈高，经理给委托人上缴的份额就愈高，故经理努力工作的积极性就愈低。当引入相对业绩比较弱化经理自己的业绩在评价企业固有生产能力上的作用（若 $\rho < \eta$），因而强化激励机制（弱化了棘轮效应）。但是，理论上谈得较多的是相对业绩比较在两个模型中的效应相反，但实际上相对业绩比较在两种情况下的效应都可能为正。因为在第一种情形，$\rho > \eta$ 是更为接近现实的假设，而在第二种情形，$\eta > \rho$ 的假定可能更为现实一些。

两类模型中绝对努力水平的差异来自我们有关直接激励的效应的假设。声誉模型中，我们曾假定显性激励合约不可行（当期的工资与当期的产出无关，经理不承担风险）；而在棘轮效应模型中，我们曾假定显性激励合约是可行的（当期工资取决于当期产出，经理承担完全的风险）。

在实际生活中，过去的业绩是可能同时都传递有两类信息的，经理既拿固定的工资（与业绩无关，但与预期能力有关），也分享剩余（这与业绩有关，但须在完成相关的目标之后）。我们可将上述结论一般化：设b为经理对自己本人的能力的所有权份额（或者是讨价还价能力），β_t为经理在t阶段的产出中占有的剩余份额。则经理在阶段 1 努力的边际收益为$(\beta_1+(b-\beta_2)\phi)$，其中$\beta_1$为直接激励，$(b-\beta_2)\phi$为隐性激励（第一阶段努力对第二阶段的收益的影响），$b\phi$度量声誉效应（强化激励），$\beta_2\phi$测量棘轮效应（弱化激励）。若$b=1$，$\beta_t=0$，则只有声誉效应，无棘轮效应，努力的边际收益为ϕ。若$b=0$，$\beta_t=1$，则仅有棘轮效应，无声誉效应，努力的边际收益为$1-\phi$（如前面的棘轮效应模型）。当引入相对业绩比较改变权数ϕ，从而既影响声誉效应也影响棘轮效应。由于相对业绩比较对这两种效应的作用恰好相反，总的激励作用不仅取决于它对ϕ的影响，而且还取决于这两种效应的相对大小。这里提到的模型未考虑显性激励机制的设计，还由于假定代理人是风险中性的，就忽略了相对业绩的比较的保险功能。但 Meyes 与 Vickers（1994）证明，在更为一般性的模型中，上述结论也是基本成立的。

9.2 锦标机制设计

9.2.1 锦标机制的特点

锦标机制（rank order tournaments）是一种常见的相对业绩比较制度。在锦标机制中，代理人的报酬多少只与他在所有代理人中的排序有关，与他的绝对努力程度不直接相关。Lazear 与 Rosen 曾证明：当代理人的业绩之间存在相关性时，即代理人们受到共同的不确定性因素影响时，通过锦标机制可以剔除更多的不确定因素，从而使委托人对代理人努力水平有更为准确的判断。这使得委托人降低风险成本，强化激励机制。此时，锦标机制是有价值的。但是，倘若代理人的业绩之间不相关，则锦标机制要比每个代理人的报酬仅依赖于他的绝对业绩水平的机制差，因为它在不增加有关代理人努力水平的信息量的情况下增大了相互竞争中的每个代理人所面临的不确定性，从而弱化激励机制。（Lazear 与 Rosen，1981）。Holmstrom（1982）证明：除非代理人面临的不确定因素是完全相关的，或者代理人的业绩只能用序数测量，否则，锦标机制不会使观测量包含的信息量得到充分利用，因为单纯的业绩排序并非反映代理人努力水平的充足统计量。他指出，当相对排序与绝对的业绩量结合起来时，委托人就可以得到一个帕累托改进。

我们在这里将要说明的是，由于委托人方也存在不兑现合约中的承诺的可能，这使得委托人方也存在一种“道德风险”，因而锦标制度在这两种情况下也有一定的意义。这种道德风险可能是源于委托人否认所观测到的业绩，且代理人也无法向委托人证实其业绩。甚至，这种信息在法庭上也无法证实。这就导致委托人在事实上观测到高产出的时候也谎称产出不高而逃避履约的责任。这种道德风险在合约规定需要向代理人支付的数额很大的情形是很可能发生的，因为存在委托人不履行合约的激励。当代理人事先就明白这一道理时，他就很可能不努力工作。但是，倘若存在一种锦标机制，它规定一定比例的代理人必将获得较高的报酬，那么，委托人的占优的选择是将较高报酬支付给业绩较高的代理人，因为这样可以激励代理人努力工作。代理人的占优战略也是选择努力工作，从而强化激励机制。我们在一个组织中所看到的“提升”制度（给定一定比例的代理人的职位会被提高），也可以用同样的道理来说明。我们在下面通过一个 Malcomson 模型来说明这一点。

9.2.2 Malcomson 模型

Malcomson（1984）假设一个企业雇佣不同年龄代的多个工人，这些工人除了在年龄代上存在差别外其它一切都是相同的。每个工人工作两个时期，因而在任何一个时期都有两代工人工作。用 $t=1,2$ 表示工龄。每个工人的效用函数为 $U(a_1,w_1,a_2,w_2)$，其中的 a_t 表示 t 时期的努力水平，w_t 表示 t 时期的工资率。自然有假定 $\frac{\partial U}{\partial a}=U_a<0$，$\frac{\partial U}{\partial w}=U_w>0$，$\frac{\partial^2 U}{\partial a^2}=U_{aa}<0$，$\frac{\partial^2 U}{\partial w^2}=U_{ww}\le 0$。设保留效用为 $\overline{U}$，劳动市场是竞争性的，从而企业可按 $EU=\overline{U}$ 的期望效用水平雇佣任何数量的工人。设每个工人在 t 时期的生产函数为

$$\pi_t=a_t+\varepsilon_t,\quad t=1,2 \qquad (9.13)$$

其中 ε_t 是均值为零的随机扰动项，可将其解释为企业对工人产出的观测误差。a_t 既为实际产出也为努力水平，π_t 为观测到的产出。

设 $F(\varepsilon_t)$ 和 $f(\varepsilon_t)$ 分别是 ε_t 的分布函数和密度函数，故给定 a_t 下 π_t 的分布函数和密度函数就分别为 $F(\pi_t-a_t)$ 和 $f(\pi_t-a_t)$。设工人存在一个必须付出的最小努力水平 $\overline{a}>0$。譬如，工人必须按时上、下班。

如果工人不能证实π_t，故根据π_t支付工资w_t的激励合约是不可行的。现规定，只要$a_t \geq \underline{a}$，就支付事前规定的标准工资。由于$U_a < 0$，故固定工资不能激励工人选择大于$\underline{a}$的努力水平。下面，我们给出基于相对业绩的锦标机制，通过它却能激励工人努力工作。

设计这样一个合约：当工人努力水平$a \geq \underline{a}$，工资率为w_L或w_H，其中$w_L < w_H$；且获得w_H的工人的比例为P。假定实际得到高工资w_H的工人的比例P是可证实的，这样的合约就是可执行的。

由于π_t与a_t是正相关的，只要得到高工资的工人是被观察到高产出的工人，这种合约就可以激励工人努力工作。这是因为企业的总的工资支出是固定的，人均期望工资等于$Pw_H + (1-P)w_L$，企业支付的工资总额为$nPw_H + n(1-P)w_L = n[Pw_H + (1-P)w_L]$，其中$n$为企业雇佣的工人总数，所以，企业的占优战略是将高工资付给业绩较好的工人。

设企业规定一个“较高业绩”的标准π^*，当$\pi_t \geq \pi^*$时，企业支付给工人高工资，其他的工人只获得低工资w_L。对于努力水平为a_t的工人，其业绩$\pi_t \geq \pi^*$的概率为$1-F(\pi^* - a_t)$。

当n充分大时，有

$$P = 1 - F(\pi^* - a_t) \tag{9.14}$$

即当企业的规模充分大时有式（9.14）。为了使合约有实际的激励作用，应有$0 < P < 1$。因为若$P = 0$，则没有人会获得高工资，$P = 1$时所有人都获得高工资。在这两种场合，都不会有激励效应。

我们以上介绍的锦标机制只在一个时期就可实施，而下面再介绍的一种类似的“提升制度”（Promotion），是在两个时期动态实施的。我们称上述合约为 1 一阶段合约，而即将给出的提升制度为 2 一阶段合约。在 2 阶段合约中，第一阶段所有工人得到相同的工资w_1；第二阶段，有P比例的工人被提升，他们得到第二阶段工资w_H，剩下的未被提升的工人得到第二阶段工资w_L。[8]

我们将证明，即使所有工人在第一阶段拿相同的工资，提升制度也能对第一阶段的努力水平产生激励效应。这一结果还可扩充到假定工人工作 T 期，如果每个阶段有 P 比例的工人被提升，则除去最后阶段后所有的前 T-1 期的工人都会努力工作。

下面，我们来考察上述 2 一阶段合约是如何激励工人在第一阶段的努力水平a_1的。

根据参与约束（见第八章），只有当期望效用不小于保留效用$\overline{U}$时，工人才会接受合约。当工人接受合约，他在阶段 1 是为w_1工作。如果企业观测到其努力带来的产出$\pi_1 \geq \pi^*$，他将被提升，时期 2 获得工资w_L，否则不被提升，时期 2 领工资w_L。

设$\overline{w}$为工人不在该企业工作的保留工资，令$w_2^* = \max\{w_L, \overline{w}\}$。则工人在决定阶段 1 的努力水平$a_1$时，将假定如果他不能被提升的话，他在阶段 2 将获得工资w_2^*。自然假定$w_H > w_2^*$，否则提升是无意义的。

假定阶段 2 后不再有提升，则$a_2 = \underline{a}$。工人的期望效用为

$$\begin{aligned} V(a_1, w_1, w_2^*, w_H, \pi^*) = F(\pi^* - a_1)U(a_1, w_1, \underline{a}, w_2^*) \\ + [1 - F(\pi^* - a_1)]U(a_1, w_1, \underline{a}, w_H) \end{aligned} \tag{9.15}$$

上式右端第一项对应工人未被提升时的情形，第二项对应工人被提升时的情形。工人在第一阶段选择最大化式（7.15）的a_1，约束条件是$a_1 \geq \underline{a}$。当存在内点解时，一阶条件为：

$$\begin{aligned} \frac{\partial V}{\partial a_1} &= -f(\pi^* - a_1)U(a_1, w_1, \underline{a}, w_2^*) + F(\pi^* - a_1) \cdot U_{a_1}(a_1, w_1, \underline{a}, w_2^*) \\ &\quad + f(\pi^* - a_1)U(a_1, w_1, \underline{a}, w_H) + [1 - F(\pi^* - a_1)] \cdot U_{a_1}(a_1, w_1, \underline{a}, w_H) \\ &= f(\pi^* - a_1)[U(a_1, w_1, \underline{a}, w_H) - U(a_1, w_1, \underline{a}, w_2^*)] \\ &\quad + F(\pi^* - a_1)U_{a_1}(a_1, w_1, \underline{a}, w_2^*) + [1 - F(\pi^* - a_1)]U_{a_1}(a_1, w_1, \underline{a}, w_H) \\ &= 0 \end{aligned} \tag{9.16}$$

当这样的解惟一时，则令其为$a_1^* = g(w_1, w_2^*, \underline{a}, w_H)$，则有

$$a_1 = \max\{a_1^*, \underline{a}\} \tag{9.17}$$

式（9.17）给出的a_1是工人在第一阶段选择的努力水平。

下面，我们来看看w_1，w_2^*，w_H和π^*怎样影响最优选择的a_1。

在式（9.14）中，合约并不直接给定π^*而是规定 P，通过该式决定π^*。[9]

我们假定，有足够多的工人，即n充分大，以致于单个工人不考虑自己的努力水平选择对π^*的影响，他在选择a_1时将π^*视为给定的。

假设效用函数U是时间可加性的，即

$$U(a_1, w_1, \underline{a}, w_H) = U'(a_1, w_1) + U^2(\underline{a}, w_H) \tag{9.18}$$

记$V_1 = \dfrac{\partial v(a_1, w_1, w_2^*, w_H, \pi^*)}{\partial a_1}$

则由式（9.16）知a_1^*满足

$$V_1 = 0 \quad （9.19）$$

据隐函数定理及式（9.19）有

$\frac{\partial a_1^*}{\partial x} = -\frac{\partial v_1}{\partial x} \Big/ \frac{\partial v_1}{\partial a_1}$，其中$x$可能是$w_1$，$w_2^*$，$w_H$和$\pi^*$中的任一个变量，我们还将$\pi^*$用 P 来替换。因$\frac{\partial U}{\partial a_1} = \frac{\partial U^1}{\partial a_1}$，$\frac{\partial U}{\partial w_1} = \frac{\partial U^1}{\partial w_1}$，据式（9.16）及（9.18），有

$$\begin{aligned} V_1 &= f(\pi^* - a_1)\left[U^2(\underline{a}, w_H) - U^2(\underline{a}, w_2^*)\right] + F(\pi^* - a_1)U'_{a_1} + \left[1 - F(\pi^* - a_1)\right]U'_{a_1} \\ &= f(\pi^* - a_1)\left[U^2(\underline{a}, w_H) - U^2(\underline{a}, w_2^*)\right] + U'_{a_1}(a_1, w_1) \\ &= f(h(P))\left[U^2(\underline{a}, w_H) - U^2(\underline{a}, w_2^*)\right] + U'_{a_1}(a_1, w_1) \end{aligned}$$

其中$h(P)$由式（9.14）决定，因为式a_1^*应满足式（9.16）。

故

$$\begin{aligned} \frac{\partial a_1^*}{\partial w_1} &= \frac{\partial g}{\partial w_1} \\ &= -\frac{\partial V_1}{\partial w_1} \Big/ \frac{\partial V_1}{\partial a_1} \\ &= -\frac{\partial U'}{\partial w_1} \Big/ \frac{\partial U'}{\partial a_1} \\ &= -\frac{U_{12}}{U_{11}} \begin{cases} \geq 0 & \mathrm{U}_{12} \geq 0 \\ < 0 & U_{12} < 0 \end{cases} \end{aligned}$$

类似有：

$$\begin{aligned} \frac{\partial a_1^*}{\partial w_2^*} &= \frac{\partial g}{\partial w_2^*} \\ &= -\frac{\partial V_1}{\partial w_2^*} \Big/ \frac{\partial V_1}{\partial a_1} \\ &= +f(a^* - a_1)\frac{U_4}{U_{11}} < 0 \end{aligned}$$

U_4在w_2^*取值。

$$\frac{\partial a_1^*}{\partial w_H}=\frac{\partial g}{\partial w_H}=-f(\pi^*-a_1)\frac{U_4}{U_{11}}>0$$，U_4在w_H取值。

$$\frac{\partial a_1^*}{\partial P}=\frac{\partial g}{\partial \pi^*}\frac{d\pi^*}{dP}=\frac{-f(\pi^*-a_1)\dfrac{d\pi^*-da_1}{dP}\left[U^2(\underline{a},w_H)-U^2(\underline{a},w_2^*)\right]}{U_{11}}$$

又由式（9.14）有

$$dP=-f(\pi^*-a_1)(d\pi^*-da_1)$$

故$\dfrac{d\pi^*-da_1}{dP}=\dfrac{1}{-f(\pi^*-a_1)}$

代入上式有

$$\frac{\partial a_1^*}{\partial P}=\frac{f'(\pi^*-a_1)\left[U(a_1,w_1,\underline{a},w_H)-U(a_1,w_1,\underline{a},w_2^*)\right]}{f(\pi^*-a_1)}$$

$$\begin{cases}\geq 0, & f'(\pi^*-a_1)\leq 0\\ <0, & f'(\pi^*-a_1)>0\end{cases}$$

这里假定了$w_H>w_2^*$，若$w_H=w_2^*$，则$\dfrac{\partial a_1^*}{\partial P}=0$。

9.2.3 Malcomson 模型的比较静态分析

当效用函数对收入和努力是可加的，则$U_{12}=0$，由上一小节知$\dfrac{\partial a_1^*}{\partial w_1}=0$，最优努力水平$a_1^*$与时期 1 的工资率$w_1$无关。这是显然的，因为$a_1^*$不影响$w_1$。

$\dfrac{\partial a_1^*}{\partial w_2^*}<0$意味着保留工资$\overline{w}$或时期 2 的低工资$w_L$的上升将导致最优努力水平下降。这是因为$w_2^*$愈高，被提升的好处就愈小，工人就愈不害怕不被提升。

$\dfrac{\partial a_1^*}{\partial w_H}>0$意味着被提升后的工资率$w_H$愈高，工人就愈努力工作，因为被提升的好处愈大。

第四个结果表明，工人对提升比例 P 的反应不是单调的。当提升比例使$f'(\pi^*-a_1)<0$，a_1^*随 P 的上升而上升，意味着提升的可能性愈大，工人工作就愈努力。但当提升比例超过某下滑界点后，由式（9.14）知π^*-a_1充分小，此时$f'(\pi^*-a_1)$一般

会大于零，故 a_1^* 随 P 的上升而下降。譬如，若 ε_t 是正态分布（均值为零），则当 $P<1/2$ 时，$f'(\pi^*-a_1)<0$，当 $P>1/2$ 时，$f'(\pi^*-a_1)>0$。此时 $P=1/2$ 就是一个临界点。其意为：

奖励面过大并不利于调动员工的积极性。当 P 太小时，满足一阶条件式（9.16）的 a_1^* 会小于 $\underline{a}$，工人也不会有积极性努力工作。

如果 $w_L \geq \overline{w}$，因而 $w_2^*=w_L$，未被提升的工人在时期 2 仍留在企业工作，若 $w_L<\overline{w}$，因而 $w_2^*=w_L$，未被提升的工人将离开企业。

企业从每个工人那里得到的期望利润为

$$\pi(w_1,w_L,w_H,P)$$
$$=\begin{cases} h(w_1,w_L,w_H,P)-w_1+\beta(\underline{a}-Pw_H-(1-P)w_L), w_L \geq \overline{w} \\ h(w_1,\overline{w},w_H,P)-w_1+\beta P(\underline{a}-w_H), w_L<\overline{w} \end{cases}$$

其中 $a_1^*=h(w_1,w_L,w_H,P)$，β 是企业的贴现因子。企业的问题是选择 (w_1,w_L,w_H,P) 最大化上述期望利润，并满足 $U \geq \overline{U}$ 和 $0 \leq P \leq 1$。

企业也可以选择 1 一阶段固定工资合约。此时，工人在二阶段的努力水平均为 $\underline{a}$，企业期望利润为

$$\pi(w_1,w_2)=\begin{cases} \underline{a}-w_1+\beta(\underline{a}-w_2), w_2 \geq \overline{w} \\ \underline{a}-w_1, \qquad\qquad w_2<\overline{w} \end{cases}$$

但这个 1 一阶段合约等价于 2 一阶段合约中 $P=0$ 的情形。因此，2 一阶段合约一定要优于 1 一阶段合约。

当我们选择 $0<P<1$ 时的期望利润大于 $P=0$ 时的期望利润，最优的 2 一阶段合约就严格优于 1 一阶段固定工资合约。

Malcomson 进一步证明（见 Malcomson,1984），只要工人第二阶段收入的边际效用 U_4 相对于第一阶段努力的边际效用的变化率 U_{11} 足够大，2 一阶段合约的期望利润就严格大于 1 一阶段固定工资合约（满足 $0<P<1$，$w_H>w_2^*$）。也就是说，从工人的最优化行为的比较静态分析看，对于任何给定的 $0<P<1$，$-U_4/U_{11}$ 决定 a_1 如何随 w_H 上升而上升；若 $-U_4/U_{11}$ 足够大，增加的 a_1（从而 π_1）就弥补额外的工资成本 $w_H-w_2^*$ 且还有余。Malcomson 在这篇论文中，用这个模型对我们在日常生活所观察到的如下 5 个方面的现象作出了解释，这 5 个现象是：

1．企业中的金字塔形等级工资制度：高工资员工的人数少于低工资员工的人数；

2．内部提升制度：有相当比例的高工资—高职位员工从内部的低工资—低职位员工中提升上来，新来的员工只在等级结构的最低层进入企业工作；

3．工龄工资：工资更多地是随工龄和经验而非生产率而上升；

4．工资差距随工龄和工作经验的增加而上升；

5．职位工资制度：工资水平将主要取决于干的是什么样的工作而非由谁来干。

Malcomson 的贡献在于用一个模型就对上述 5 个现象作出了统一的解释，而在他之前的工作仅能分别解释其中的 1 到 2 个现象。

注释

[1] 注意，正是这一点，张五常在其《佃农理论》中提出地主无需直接观测佃农的努力水平，通过观测竞争性的佃农市场中的众多变量就可推断出土地的合理租金率。

[2] 国有企业的产出除了利润外还包括一些“社会效益”，如提供就业机会，体现“社会公平”等，但由于这类“社会效益”难以定量计测，经理与其政府管理部门之间在“产出”的大小上难以取得一致。所以，对国有企业经理一般没有显性合约的激励。正是由于此点原因，国有企业经理的个人货币收入都较低（作为企业“产出”的一部分即“社会效益”没有成为经理货币收入的一个贡献因素）。所以，国有企业经理一般都主要关心自己的仕途而非货币收入，因为仕途上的升迁不必按定量化的显性激励合约决定。当然，至少在中国大陆，一些国有企业近来已在进行“改革”，试图通过年薪制将显性激励机制引入国有企业，但正如许多研究所指出的那样，这种努力的最终效果是值得怀疑的。

[3] 在竞争性经理人市场上，经理的工资率水平由边际生产率决定。而经理的边际生产率据我们这里的假设等于$\pi_t = a_t + \theta + u_t$。在第二阶段工资率水平由$\pi_2$的期望值决定，它与努力水平及能力水平都有关。所以，如果经理在第一阶段努力工作有助于提高第二阶段经理市场或股东对其能力水平的评价，那么，经理人就可能在第一阶段努力工作，即使在第一阶段没有任何显性激励机制。

[4] 注意到式（9.6）中有一个随机变量u_1，即经理人的效用是一个随机变量。按照习惯，在取得最优化一阶条件时应该先计算其期望效用，然后对期望效用取极大值。由于这里假定经理人是风险中性的，期望效用最大化等价于效用最大化，故这里没有根据期望效用导出一阶条件，而是直接计算效用函数最大化的一阶条件。在这里，无论随机变量u_1取何值，最优化一阶条件都是相同的，即a_1和a_2在与u_1无关的位置上使U达到最大，因此有这里的数学处理。

[5] 如果τ具有最大的传递θ的信息的功能，即$\tau = 1$，则第一阶段帕累托最优努力水平就能够实现，尽管这里没有显性激励机制。

[6] 这里假定θ与时间有关，与前面的简化假定是不同的。

[7] 按理说这里还有一个代理人参予约束的问题，故是一个约束极大化的问题。这里不妨设参予约束总成立，即不考虑这一问题。同时，根据代理人风险中性假定，期望效用最大化等价于效用最大化，这与前一小节的分析类似。

[8] 表面看来 2 一阶段合约与 1 一阶段合约起到的预期作用是相同的，但实际上 2 一阶段合约更加可行一些。这是因为，1 一阶段合约要求工资的支付在期末进行，因为业绩评估有成本，这样的评估只能有限次地进行，但工人要在期间消费。企业进入资本市场的成本要低一些，故企业预付工资更加有效一些。但预付工资不能依赖于本期的业绩。在 2 一阶段合约中，本期的业绩只影响下一期的工资，企业无须等到期末再支付工资。当然，因为工人在第二阶段后就要退休，没有继续提升的可能，这样的 2 一阶段合约就不能提供第二阶段的激励，工人将选择$a_1 = \underline{a}$。一种方法是根据第二阶段的表现发放不同的退休金，即一定比例的工人得到高的退休金，其他工人得到低退休金。

[9] 因为给定了P，就能将π^*定为排名为$Pn+1$的工人的业绩。

第十章　多任务代理及产权：委托—代理模型的一种扩充

如果我们假定代理人实际上为委托人代理了不止一项的多项任务，或者说代理人的工作有多个维度或目标，则我们在前面各章中获得的委托—代理理论就可能是不适用的了。在前面的第八、九章中，我们潜在地假定了代理人为委托人代理了仅仅是一项工作或代理人的工作只有一个维度，并假定委托人不能（经济地）观测到代理人的努力程度，这样，作为委托人的利益使然，机制设计要求代理人获得的报酬必须依赖于可观察到的业绩，这样才可能激励代理人为委托人努力工作。

现在，我们将代理人的工作目标只有一个维度的假定放松，假定代理人需要同时完成多项任务或进行多个维度的工作。我们来看看在这种情形下最优的委托—代理合约是什么？什么是代理人的多维度工作或多种任务呢。工人的业绩既包括其生产的数量，又包括产品的质量；企业经理不仅要完成短期的利润，而且还应考虑企业资产的长期增值和企业未来的长期盈利能力；销售经理不仅要通过拉客户增加销售量，而且还应搞好售后服务，从而增强企业的长期销售量，教师不仅要向学生教授书本知识，而且还应注重培养学生的创造力和想象力等素质等等。由于时间是稀缺的资源，代理人在多项任务上的时间分配就存在此消彼长的效应。譬如，企业经理注重短期利润时，就可能通过拼设备而损害企业未来的长期生产能力。

现在的问题是，当代理人的任务有多项时，委托人往往对不同工作的监督能力是不同的，有一些工作可能比另一些工作更加难以监督。譬如，工人生产的产品数量显然比产品质量更容易监督一些；观测企业的当期利润比测度企业的资产价值及未来盈利能力要容易多了；当代理人是销售人员时，易于监督的是代理人直接花在销售活动上的努力，因为这种努力会直接增加可观测到的销售量，而其它的活动只增加后来一段时间后的销售量；如果代理人是教师，委托人可通过学生的考试成绩判断学生学到的知识（不完全准确），但要测度学生的创造力和想象力从而判断教师在这方面花费了多少精力是十分难的。

当对不同工作的监督有着不同的难易程度时，对易于监督的工作的过度激励会诱使代理人将过多的努力花在这些方面而忽视其它方面，从而扭曲资源配置。譬如，计件工资制虽可以调动工人增加产品数量的积极性，但却不利于维持和改进产品的质量（当质量不易监督时）；若将经理的报酬过度地依赖于利润，则可能会诱使经理只注重短期的利润而不注重企业的长期资产增值；只根据销售量对销售人员进行奖惩会使销售人员忽视售后服务等其它有益的活动；将教师的报酬与学生的考试成绩或升学率挂钩可能诱使教师不注意培养学生的创造力和想象力，这是应试教育的弊病，不利于素质教育。

Holmstrom 与 Milgrom（1991）证明了这一直觉，即当代理人从事多项工作时，从简单的委托—代理理论得出的结论可能是不恰当的。他俩发现，在一些情况下，固定工

资合约反而可能优于根据可观测变量奖惩代理人的激励合约。事实上，他俩在这篇论文中还对如下一些现象作了解释，即（1）为什么资产的所有权是重要的；（2）为什么有时候雇佣制度优于独立合约制度即市场交易制度，即使前者本身并没有专业化的优势，并且资本市场是完全的；（3）为什么政府官员行动的自由度受到更多的限制；（4）为什么有时候让一个人专业化于一项工作可能更加有效，即使专业化本身并没有技术上的优势。

在本章中，我们介绍该模型的基本内容。

10.1 Holmstrom-Milgrom 多任务模型

假定代理人从事两项工作，努力水平的选择是同时进行的，由此构成一个静态博弈。设表示代理人努力程度的向量是 $a=(a_1,a_2)$，其中 a_1 是花在第一项工作上的努力，a_2 是花在第二项工作上的努力；用 $B(a_1,a_2)$ 表示努力的期望收益（所有权属于委托人），用 $C(a_1,a_2)$ 表示努力的成本（代理人直接承受的），并设 $B(a_1,a_2)$ 是严格递增的凹函数，$C(a_1,a_2)$ 是严格递增的凸函数。

当代理人努力选择为 $a=(a_1,a_2)$ 时，决定了一个可观测的信息向量 X，并有

$$X=\mu(a_1,a_2)+\varepsilon$$

假定 μ 是从 R_+^2 映射到 R^k 的凹函数，其中 R 是实数域，R^k 是 k 维欧氏空间 $(k\geq 0)$，k 是可观测信息的数量。这样，两个努力变量 a_1 和 a_2 决定了 k 个可观测信息。ε 是服从正态分布的随机向量，设其值为 0，协方差矩阵为 Σ'。此时，X 就是服从均值为 $\mu(a_1,a_2)$，协方差矩阵为 Σ 的正态分布。

当 $k=2$ 时，一种特别的情形是

$$X=\begin{bmatrix}x_1\\x_2\end{bmatrix}$$

$$x_1=\mu_1(a_1)+\varepsilon_1$$

$$x_2=\mu_2(a_2)+\varepsilon_2$$

$$\mu=\begin{bmatrix}u_1\\u_2\end{bmatrix}$$

此时不同的努力变量产生了不同的信息。如果 ε_1 与 ε_2 相关，则不同的信息仍然是相关的。这种情形里，x_1 就反映了 a_1，x_2 反映了 a_2。譬如，当 a_1 是刻画代理人花在“数量”上的努力，a_2 刻画代理人花在“质量”上的努力时，x_1 就可以被理解为是观测到的“数量”，x_2 就是观测到的“质量”。

假定委托人是风险中性的，代理人是风险规避的。并进一步假设代理人具有不变的绝对风险规避度，努力成本 $C(a_1,a_2)$ 是用货币等价物来测定的。则当代理人的工资函数为 $S(X)$，假定其为线性函数 $S(X)=\alpha+\beta^T X$，其中 $\beta^T=(\beta_1,\cdots,\beta_k)$，$T$ 表示向量的转

置。由于代理人的收入$S(X)$是随机变量，我们来计算他的确定性等价收入。

在第8章中，我们定义$W=S(X)-C(a_1,a_2)$的确定性等价收入为满足$u(CE)=Eu(w)$的 CE，其中u为代理人的效用函数，E 表示数学期望。类似于第 8 章中的处理，这里也假定代理人的效用函数为$u(w)=-e^{\rho_w}$。由此有$-\exp[-\rho CE]=\int[-\exp(-\rho\omega)f(\omega)dw$，其中$f(w)$是 S 的密度函数。

故

$$
\begin{aligned}
-\exp[-\rho CE] &= -\int\exp\{-\rho[S(X)-C(a_1,a_2)]\}f(w)dw \\
&= -\int\exp\{-\rho[\alpha+\beta^T X-C(a_1,a_2)]\}f(w)dw \\
&= -\int\exp\{-\rho[\alpha+\beta^T(\mu(a_1,a_2)+\varepsilon]-C(a_1,a_2)]\}f(\varepsilon)d\varepsilon \\
&= -\int\exp\{-\rho[\alpha+\beta^T\mu(a_1,a_2)-C(a_1,a_2)+\beta^T\varepsilon]\}f(\varepsilon)d\varepsilon \\
&= -\int\exp\{-\rho[\alpha+\beta^T\mu(a_1,a_2)-C(a_1,a_2)]\}\exp[-\rho\beta^T\varepsilon]f(\varepsilon)d\varepsilon
\end{aligned}
$$

因为 $f(\varepsilon)=A\exp\left(-\frac{\varepsilon^T\Sigma^{-1}\varepsilon}{2}\right)$，其中 $A=(2\pi)^{\frac{k}{2}}|\Sigma|^{-\frac{1}{2}}$为常数。

故

$$
\begin{aligned}
-\exp[-\rho CE] &= -A\exp\{-\rho[\alpha+\beta^T u(a_1,a_2)-C(a_1,a_2)]\} \\
&\quad\cdot\int\exp\left[-\rho\beta^T\varepsilon-\frac{\varepsilon^T\Sigma^{-1}\varepsilon}{2}\right]d\varepsilon \\
&= -A\exp\{-\rho[\alpha+\beta^T u(a_1,a_2)-C(a_1,a_2)]\} \\
&\quad\cdot\int\exp\left[-\rho\beta^T\varepsilon-\frac{\varepsilon^T\Sigma^{-1}\varepsilon}{2}-\frac{\rho^2\beta^T\Sigma\beta}{2}+\frac{\rho^2\beta^T\Sigma\beta}{2}\right]d\varepsilon \\
&= -A\exp\{-\rho[\alpha+\beta^T u(a_1,a_2)-C(a_1,a_2)]\}\exp\left(\frac{\rho^2\beta^T\Sigma\beta}{2}\right) \\
&\quad\cdot\int\exp\left[-\frac{(\varepsilon+\rho\Sigma\beta)^T\Sigma^{-1}(\varepsilon+\rho\Sigma\beta)}{2}\right]d\varepsilon \\
&= -\exp\left\{-\rho[\alpha+\beta^T u(a_1,a_2)-C(a_1,a_2)]-\frac{\rho^2\beta^T\Sigma\beta}{2}]\right\} \\
&\quad\cdot(2\pi)^{-\frac{k}{2}}|\Sigma|^{-\frac{1}{2}}\int\exp\left[-\frac{\tilde{\varepsilon}^T\Sigma^{-1}\tilde{\varepsilon}}{2}\right]d\varepsilon
\end{aligned}
$$

其中$\tilde{\varepsilon}=\varepsilon+\rho\Sigma\beta$，根据分布函数的归一化性质，有

$$(2\pi)^{-\frac{k}{2}}|\Sigma|^{-\frac{1}{2}}\int\exp\left[-\frac{\tilde{\varepsilon}^T\Sigma^{-1}\tilde{\varepsilon}}{2}\right]d\tilde{\varepsilon}=1$$

于是：

$$-\exp[-\rho CE]=-\exp\left\{-\rho[\alpha+\beta^T u(a_1,a_2)-C(a_1,a_2)]-\frac{\rho\beta^T\Sigma\beta}{2}]\right\}$$

即 $CE=\alpha+\beta^T\mu(a_1,a_2)-\frac{\rho\beta^T\Sigma\beta}{2}--C(a_1,a_2)$

其中 $\alpha+\beta^T\mu(a)$ 是期望工资，ρ 为绝对风险规避度，$\beta^T\Sigma\beta$ 是代理人的收入方差。根据上式，$\frac{1}{2}\rho\beta^T\Sigma\beta$ 就是风险贴水或委托人的风险成本。

委托人的期望利润是：

$$B(a_1,a_2)-E[S(X)]=\beta(a_1,a_2)-\alpha-\beta^T\mu(a_1,a_2) \qquad (10.1)$$

固定工资 α 由代理人的保留支付 $\overline{u}$ 决定（参与约束）。

由于委托人是风险中性的，故其期望利润就是其确定性等价收入。代理人的激励相容约束为

$(a_1,a_2)\in\arg\max CE$

它等价于 $(a_1,a_2)\in\arg\max CE[\beta^T u(a_1,a_2)-C(a_1,a_2)]$ （10.2）

参与约束为

$$CE=\alpha+\beta^T u(a_1,a_2)-\frac{\rho\beta^T\Sigma\beta}{2}-C(a_1,a_2)\geq\overline{u} \qquad (10.3)$$

委托人的问题是选择 β 使其确定性等价收入在参与约束和激励相容约束下达到最大化。显然，在最大化委托人确定性等价收入情况下，参与约束（10.3）必为等式，于是有

$$\beta^T u(a_1,a_2)=\overline{u}-\alpha+\frac{\rho\beta^T\Sigma\beta}{2}+C(a_1,a_2)$$

代入式（10.1），委托人的确定性等价收入为

$$B(a_1,a_2)-u-\frac{\rho\beta^T\Sigma\beta}{2}-C(a_1,a_2)$$

这样，委托人的问题就变成如何选择 β 使总的确定性等价收入

$$[B(a_1,a_2)-\alpha-\beta^T u(a_1,a_2)]+[\alpha+\beta^T\mu(a_1,a_2)-\frac{1}{2}\rho\beta^T\Sigma\beta-C(a_1,a_2)]$$

$$=B(a_1,a_2)-\frac{1}{2}\rho\beta^T\Sigma\beta-C(a_1,a_2)$$

最大化。

下面，为了得到具体的结果，我们再进一步假定 $u(a_1,a_2)=(a_1,a_2)^T$，观测变量就为 $x_i=a_i+\varepsilon_i$，$i=1,2$。不妨设有 $a_i>>0$，$i=1,2$，则激励相容约束（10.2）变为

$$\beta_i=\frac{\partial C(a_1,a_2)}{\partial a_i}=C_i(a_1,a_2),\quad i=1,2 \qquad (10.4)$$

式（10.4）决定了努力函数 $a_i=a_i(\beta)$。

在式（10.4）两端对 a_i 求导

$$\frac{\partial \beta}{\partial a} = [a_{ij}] \qquad （10.5）$$

在式（10.4）两端对 β 求导，得

$$\left[\delta_{ij}\right] = \left[C_{ij}\right]\frac{\partial a}{\partial \beta} \qquad （10.6）$$

其中 $\delta_{ij} = \begin{cases} 1 & i \neq j \\ 0 & i \neq j \end{cases} \quad i = 1,2$

故

$$\frac{\partial a}{\partial \beta} = \left[C_{ij}\right]^{-1} \qquad （10.7）$$

这里有

$$\frac{\partial B}{\partial a} = \begin{bmatrix} \dfrac{\partial B_1}{\partial a_1} & \dfrac{\partial B}{\partial a_2} \\ \dfrac{\partial B_2}{\partial a_1} & \dfrac{\partial B_2}{\partial a_2} \end{bmatrix}，\quad \left[C_{ij}\right] = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}$$

$$C_{ij} = \frac{\partial C_i}{\partial a_j}，\quad i = 1,2$$

最大化总的确定性等价收入的一阶条件为：

$$\left(\frac{\partial B}{\partial a}\right)^T \frac{\partial a}{\partial \beta} - \rho^2 \beta^T \Sigma - \left(\frac{\partial C}{\partial a}\right)^T \frac{\partial a}{\partial \beta} = 0$$

由激励相容约束（10.4）有 $\left(\dfrac{\partial C}{\partial a}\right) = \beta$ 。

故有

$$\left(\frac{\partial B}{\partial a}\right)^T \frac{\partial a}{\partial \beta} - \rho \beta^T \Sigma - \beta^T \frac{\partial a}{\partial \beta} = 0$$

$$\beta^T = \left(\frac{\partial \beta}{\partial a}\right)^T \frac{\partial a}{\partial \beta}\left[\rho\Sigma + \frac{\partial a}{\partial \beta}\right]^{-1}$$

$$= \left(\frac{\partial \beta}{\partial a}\right)^T \frac{\partial a}{\partial \beta}\left[\left(\rho\Sigma + \left(\frac{\partial a}{\partial \beta}\right)^{-1} + I\right)\frac{\partial a}{\partial \beta}\right]^{-1}$$

$$= \left(\frac{\partial \beta}{\partial a}\right)^T \frac{\partial a}{\partial \beta}\left(\frac{\partial a}{\partial \beta}\right)^{-1}\left[I + \rho\Sigma\left(\frac{\partial a}{\partial \beta}\right)^{-1}\right]^{-1}$$

$$= \left(\frac{\partial \beta}{\partial a}\right)^T \left[I + \rho\Sigma\left(\frac{\partial a}{\partial \beta}\right)^{-1}\right]^{-1}$$

其中 I 为单位矩阵，由式（10.7），有 $\beta^T = \left(\frac{\partial B}{\partial a}\right)^T \left[I + \rho\Sigma\left(C_{ij}\right)\right]^{-1}$。

故[1]

$$\beta = \left[I + \rho\left(C_{ij}\right)\Sigma\right]^{-1} B' \qquad (10.8)$$

其中，$B' = \frac{\partial B}{\partial a}$

注意，在我们的假定下，式（10.8）是最大化的充分必要条件。

当 a 和 X 都是一维变量时，式（10.8）变为

$$\beta = \frac{B'}{1 + \rho \frac{\partial^2 C}{\partial a^2}\sigma^2}$$

令 $C = ba^2/2$，$B = a$，则 $B' = 1$，$\frac{\partial^2 C}{\partial a^2} = b$，则

$$\beta = \frac{1}{1 + \rho b \sigma^2}$$

这正是我们在第 8 章中得到的最优合约。

下面，我们看看当 ε 中各分量是独立同分布的随机变量和两种努力的成本函数是独立的时，最优合约具有什么样的特征。此时，Σ 是对角矩阵，而 $C_{ij} = 0$，$i \neq j$。由式（10.8）有

$$\beta_i = \frac{B_i}{1 + \rho C_{ii}\sigma_i^2}，\quad B_i = \frac{\partial B}{\partial a_i}，\quad i = 1,2$$

此时，如果 $B_i = B_i(a_i)$，$i = 1,2$，即每一种工作的努力不影响其它工作努力程度的生产率，以最优的 β 是相互独立的，因为代理人在给定工作上的最优努力独立于在其它工作上的努力程度。注意到此时 β_i 是 ρ，C_{ii} 和 σ_i^2 的减函数，这与第 8 章中得到的结果是相同的。

10.2 模型分析与固定工资合约

观察一下式（10.8），一般地，C_{ij} 当 $i \neq j$ 时是不为零的。C_{ij} 不为零说明代理人的努力成本在不同工作中存在相互依存性。式（10.8）说明，这种相互依存性在决定最优工资合约 $S(X)$ 上有着重要的作用。下面，我们对此展开讨论。

假定代理人是教师，a_1 为他花在基本知识教育上的努力，a_2 为他花在培养学生创造力和想象力的努力。我们假定 a_2 是不可测度的。唯一的信息变量是[2]

$$x_1 = a_1 + \varepsilon$$

假定有 $a >> 0$，则由式（10.8）有

$$\beta_1=\left[I+\rho[C_{ij}]\Sigma\right]^{-1}B'$$

$$=\left[I+\rho\begin{pmatrix}C_{11}&C_{12}\\C_{21}&C_{22}\end{pmatrix}\begin{pmatrix}\sigma_1^2&\sigma_{12}\\\sigma_{21}&\sigma_2^2\end{pmatrix}\right]^{-1}B'$$

$$=\left[I+\rho\begin{pmatrix}C_{11}\sigma_1^2+C_{12}\sigma_{21}&C_{11}\sigma_{12}+C_{12}\sigma_2^2\\C_{21}\sigma_1^2+C_{22}\sigma_{21}&C_{21}\sigma_{12}+C_{22}\sigma_2^2\end{pmatrix}\right]^{-1}B'$$

$$=\left[I+\rho\begin{pmatrix}C_{11}\sigma_1^2&C_{12}\sigma_2^2\\C_{21}\sigma_1^2&C_{22}\sigma_2^2\end{pmatrix}\right]^{-1}B'$$

$$=\begin{bmatrix}1+\rho C_{11}\sigma_1^2&\rho C_{12}\sigma_2^2\\\rho C_{21}\sigma_1^2&1+\rho C_{22}\sigma_2^2\end{bmatrix}^{-1}B'$$

$$=D^{-1}B'$$

其中，$D=\begin{bmatrix}1+\rho C_{11}\sigma_1^2&\rho C_{12}\sigma_2^2\\\rho C_{21}\sigma_1^2&1+\rho C_{22}\sigma_2^2\end{bmatrix}$。

下面计算D^{-1}，根据基本的代数知识，我们有$D^{-1}=\dfrac{D^*}{\det D}$，其中矩阵$D^*$中第$i$行$j$列上的元素是$D$中第$j$行$i$列元素的代数余子式，$\det D$为$D$的行列式。

故$D^*=\begin{bmatrix}1+\rho C_{22}\sigma_2^2&-\rho C_{12}\sigma_2^2\\-\rho C_{21}\sigma_1^2&1+\rho C_{11}\sigma_1^2\end{bmatrix}$

$$\det D=(1+\rho C_{11}\sigma_1^2)(1+\rho C_{22}\sigma_2^2)-\rho^2C_{12}^2\sigma_1^2\sigma_2^2$$

（利用$C_{12}=C_{21}$）$=1+\rho C_{22}\sigma_2^2+\rho C_{11}\sigma_1^2+\rho^2C_{11}C_{22}\sigma_1^2\sigma_2^2-\rho^2C_{12}^2\sigma_1^2\sigma_2^2$

$$=1+\rho[C_{11}\sigma_1^2+C_{22}\sigma_2^2)+\rho^2[C_{11}C_{22}-C_{12}^2)\sigma_1^2\sigma_2^2$$

$$=d$$

所以$D^*=\begin{bmatrix}\dfrac{1+\rho C_{22}\sigma_2^2}{d}&\dfrac{-\rho C_{12}\sigma_2^2}{d}\\\dfrac{-\rho C_{21}\sigma_1^2}{d}&\dfrac{1+\rho C_{11}\sigma_1^2}{d}\end{bmatrix}$

令$\sigma_2^2\to+\infty$，则

$$D^*=\begin{bmatrix}\dfrac{\rho C_{22}}{\rho C_{22}+\rho^2(C_{11}C_{22}-C_{12}^2)\sigma_1^2}&\dfrac{-\rho C_{12}}{\rho C_{22}+\rho^2(C_{11}C_{22}-C_{12}^2)\sigma_1^2}\\0&0\end{bmatrix}$$

得到

$$\beta_1=\frac{B_1-B_2C_{12}/C_{22}}{1+\rho\sigma_1^2(C_{11}-C_{12}^2/C_{22})}\qquad（10.9）$$

当$C_{12}<0$时，$|C_{12}|$愈大，β_1就愈大。如果设C_{ij}为常数，则β_1增大时，a_1和a_2都可

能增大，这说明当教师花在基本知识教育上的努力与花在创造性教育上的活动在成本上是互补的时，对基本知识教育活动的激励就应该加强。相反，若$C_{12}>0$，即两种活动在成本上是替代的时，如果设C_{ij}为常数，则当β_1增大时，可能使a_1增大而a_2减小。此时，对基本知识教育活动的激励就应该弱化，因为较高的β_1将诱使教师把过多的精力花在基本知识教育上而忽视对学生创造力和想象力的培养。一般地，总有$C_{12}>0$，即代理人的精力是有限的，一种工作愈努力，则另一种工作的边际成本就愈高。此时，委托人有两种办法诱使代理人增加在任何给定活动上的努力水平。即或者直接奖励该种活动，或者减少该活动的机会成本（即弱化对其它活动的激励）。上面我们假定a_2不可观测，故激励教师在培养学生创造力方面花精力的唯一办法就是减少β_1，即弱化对基本知识教育的激励。

根据式（10.9），在一定的条件下，$\beta_1=0$可能是最优的合约。譬如，只有当$a_1>0$，$a_2>0$时，才有$B>0$，此时$B_1(a_1,0)=B_2(0,a_2)=0$，因$\sigma_2^2=+\infty$的假设导致$\beta_2=0$，当假设$C_2(a_1,0)=0$时，由$\beta=\dfrac{\partial C}{\partial a}$知$\beta_2=C_2(a_1,a_2)$，即有$a_2=0$，故$B=0$，此时激励代理人在第一种工作上的努力不能给委托人带来收益，故$\beta_1=0$是最优的。这说明当我们并不单纯追求“数量”或对学生培养中不片面追求知识掌握时（即$B>0$当且仅当$a_1a_2\neq 0$），则$\beta_1=0$是最优的（假定$C_2(a_1,0)=0$）。

由于$\beta_1=\beta_2=0$，若假设$C_1(0,a_2)=C_2(a_1,0)=0$，则由$\beta=\dfrac{\partial C}{\partial a}$得$a_1=a_2=0$，总的确定性等价收入为

$$B(0,0)-\frac{1}{2}\rho\begin{pmatrix}0\\0\end{pmatrix}^T\Sigma\begin{pmatrix}0\\0\end{pmatrix}-C(0,0)=0$$

但当$\beta_1>0$，$a_1>0$时，总的确定性等价收入为

$$\begin{aligned}&B(a_1,0)-\frac{1}{2}\rho\begin{pmatrix}\beta_1\\0\end{pmatrix}^T\Sigma\begin{pmatrix}\beta_1\\0\end{pmatrix}-C(a_1,0)\\&=-\frac{1}{2}\rho(\beta_1,0)\begin{pmatrix}\sigma_1^2&\sigma_{12}\\\sigma_{21}&\sigma_{22}\end{pmatrix}\begin{pmatrix}\beta_1\\0\end{pmatrix}-C(a_1,0)\\&=-\frac{1}{2}\rho\beta_1^2\sigma_1^2-C(a_1,0)<0\end{aligned}$$

这说明对第一种工作的单独的激励只会减少总的确定性等价收入。在每种努力有独立的价值时，如$B(a_1,a_2)=B^1(a_1)+B^2(a_2)$，此时才有必要对第一种工作加以单独的激励。

有一种特别的情形，$C(a_1,a_2)=C(a_1+a_2)$，此时两种努力在成本上是完全替代的。由$\beta=\dfrac{\partial C}{\partial a}$知$\beta_1=\beta_2$（代理人在不同工作上的边际收益相等）[3]。因为当$\beta_1\neq\beta_2$时，代

理人不会在边际收益低的工作上作任何努力。如当$\beta_1 > \beta_2$时，有$a_2 = 0$。

如果$\sigma_2^2 = 0$，则$\beta_1 = \beta_2 = 0$。

上面的分析使我们得到一个可能成立的结论，即当代理人从事多任务工作时，对给定工作的激励不仅取决于该工作中的可观测性，而且取决于其它工作的可观测性。如果委托人期待代理人在一项不可观测的工作上多加努力，则激励工资不应用于任何的其它工作。

根据这种理论，我们可以解释现实中存在的一些制度安排。

固定工资制度。在第 8 章中，我们获得这样一种推论，即当委托人不能观测代理人的努力但可以观测到与代理人努力相关的变量时，根据观测到的变量向代理人支付收入的激励合约是优于固定工资合约的。但是，尽管可观测变量在现实中一般都是存在的，我们却普遍观测到固定工资合约。这种现象似乎可用我们在上面给出的模型加以说明[4]。

当工作的努力不仅决定产品数量，而且还决定产品质量时，当质量是重要的且又是难以监督的，则计时工资（固定工资）就可能是优于计件工资（激励工资）的制度。对于教师的工资来说更是如此。

为了使这种解释成立，还要假定在固定工资下，代理人会选择严格正的努力水平才行。这是一个现实必须的假设，也当然是成立的，因为委托人有办法监督代理人完成一个起码的工作量，或者还有一种可能，即当努力水平小于某个水平时，工作并不带来负的效用。或者这两种因素都发生作用。譬如，委托人也容易对代理人是否上班进行监督，而对代理人来说，给定自己已经在班上，不干任何事情是比干点事更加痛苦的，而当需求代理人更进一步努力工作时，才需要激励。

下面，我们用一个例子来说明上面的模型。

例 10.1　设两种活动在成本函数上是完全替代的，即

$C(a_1, a_2) = C(a_1 + a_2)$

令$\underline{a} > 0$为固定工资制度下代理人的努力水平，满足$C(\underline{a}) = 0$，进一步特别假定成本函数就是

$C(a_1 + a_2) = (a_1 + a_2 - \underline{a})^2 / 2$

收益函数为$B(a_1, a_2) = \sqrt{a_1, a_2}$

第一种工作（如产品数量）是可观测的，$0 < \sigma_1^2 < \infty$，第二种工作（如产品质量）是不可观测的，$\sigma_2^2 = \infty$。

在固定工资制度下，总的努力$\underline{a}$的最优分配为$a_1 = a_2 = \underline{a}/2$，即$a_1 - \dfrac{\underline{a}}{2} = a_2 - \dfrac{\underline{a}}{2} = 0$。由于总的努力分配方式并不影响代理人的效用，故假定代理人将根据委托人的意志选择$a_1 = a_2 = \underline{a}/2$。总的确定性等价收入为

$$B(\underline{a}/2, \underline{a}/2) - C(\underline{a}) = \sqrt{\left(\frac{\underline{a}}{2}\right)^2} - C(\underline{a}) = \frac{\underline{a}}{2}$$

代理人的收入为 $S(X) = \alpha + \beta^T X$，此时只有一个可观测量 x_1，故 $S(x_1) = \alpha + \beta_1 x_1$，其确定性等价收入为

$$CE = \alpha + \beta_1 a_1 - \frac{\rho}{2}\beta_1^2 \sigma_1^2 - C(a)$$

当 $\beta_1 > 0$ 时，CE 的最大化条件必为 $a_2 = 0$（是边界点），而 a_1 满足

$$\frac{\partial CE}{\partial a_1} = \beta_1 - (a_1 + a_2 - \underline{a}) = 0$$

即 $a_1 = \beta_1 + \underline{a} > \underline{a}$

于是对于任何 $\beta_1 > 0$，有 $a_1 > \underline{a}$。这说明将代理人的收入与观测到的变量 $x_1 = a_1 + \varepsilon_1$ 挂钩可激励代理人更加努力地从事第一项工作，但激励工资又诱使代理人将原来用于第二种工作的努力 $\underline{a}/2$ 转移到第一种工作上。结果此时的总确定性等价收入为

$$B(\underline{a} + \beta_1, 0) - \frac{1}{2}\rho\beta^2\sigma_1^2 - C(\beta_1)$$

$$= -\frac{1}{2}\rho\beta_1^2\sigma_1^2 - \frac{1}{2}\beta_1^2$$

$$= -\frac{1}{2}\beta^2(\rho\sigma_1^2 + 1) < 0$$

此时固定工资合约显然要优于激励合约。

我们可用图 10.1 来说明上述分析。

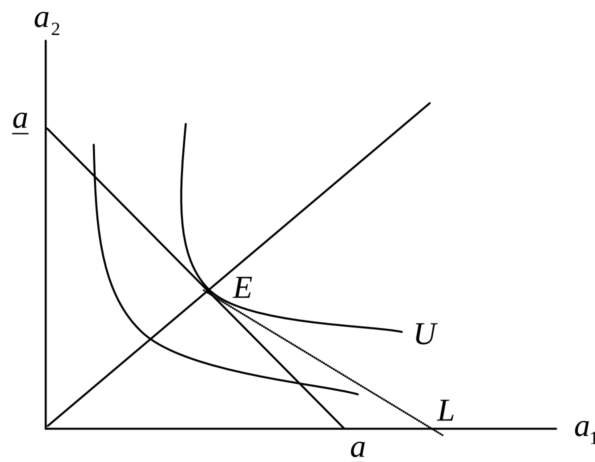


图 10.1　成本可替代情形的最优努力组合

在图 10.1 中，$\underline{a}\underline{a}$ 线是成本完全可替代时的代理人两种工作努力组合线。U 是总确定性收入的等高线或无关差异曲线。横轴是对第一种工作的激励线。显然，U 与 $\underline{a}\underline{a}$ 线

的切点 E 是最优努力组合，它总是优于仅对第一种工作进行激励的情形。图中的虚线 L 代表代理人两种工作努力成本不完全可替代的情形。即使在这种情形，激励仍然劣于固定工资的情形。

作为这个模型的一个应用，我们将第一种工作解释为一个组织中代理人个人负责的工作，第二种工作解释为几个代理人集体负责的工作，当我们假定个人对集体的贡献不可度量时，则当委托人想使个人有兴趣对集体作贡献时，固定工资就是最优的，这说明大锅饭制度在一定目标下也有其存在的合理性呢！

10.3 产权制度安排

应用上一小节中的代理人多任务模型的思路，我们可以研究这样的一个问题，即倘若委托人与代理人在合约的执行过程中涉及到资产的使用，那么，资产的所有权最好应安排给谁最好呢？譬如，当出租车公司与出租车司机签订合约时，出租车最好是由公司提供呢还是由司机自己提供？在这里，我们将资产的产权定义为资产所有者在合约结束后重新安排资产用途的权利。当出租车属于公司时，在司机与公司的合约结束后，公司可以将车继续交给原来的司机使用，也可以将车交给另一个司机使用，或将车卖掉。如果出租车的所有权归司机，则合约结束后，司机可与原来的公司续签合约，也可以自己带着车与另一个公司签约，或将车卖掉。在下面的分析中，我们假设代理人从事的第一项工作是使用资产创造利润，第二项工作是维护和保养资产。假定期望收益函数为 $B(a_1)+V(a_2)$，其中 $B(a_1)$ 是期望利润，$V(a_2)$ 是期望资产价值；可观测的实际利润为 $B(a_1)+\varepsilon_b$，实际的资产价值为 $V(a_2)+\varepsilon_v$，其中 ε_b 和 ε_v 是服从正态分布的随机变量，它们的均值为 0，方差分别为 σ_b^2 和 σ_v^2，协方差设为 0（实际利润的扰动因素与实际资产的扰动不相关）。设利润 $B(a_1)+\varepsilon_b$ 是归委托人所有的，资产的价值 $V(a_2)+\varepsilon_v$ 归资产的所有者所有。设 $B(a_1)$ 和 $V(a_2)$ 都是严格递增的凹函数，且 a_1 和 a_2 在效用函数上是完全可替代的，最小努力水平为 $\underline{a}$。

在现实中，资产的价值是很难准确评估的。于是假定：尽管代理人花在维护和保养资产上的努力 a_2 影响资产价值 $V(a_2)+\varepsilon_v$，但委托人没有可观测的信息来评价 a_2。所以，委托人可获得的唯一可观测变量是实际利润 $x=B(a_1)+\varepsilon_b$。

如果资产属于委托人所有，激励合约为 $S(x)=\alpha+\beta x$。给定 a_1 和 a_2 在效用函数上是完全可替代的，为了使代理人在资产维护和保养上也花一定精力，则委托人的最优选择是 $\beta=0$，即不给予代理人任何的激励。此时，代理人将 $\underline{a}$ 最优地分配于两种工作（我们分别用 $\underline{a_1}$ 和 $\underline{a_2}$ 来表示），总的确定等价收入为[5]：

$$B(\underline{a_1})+V(\underline{a}-\underline{a_1})-\frac{1}{2}\rho Var(0)-C(\underline{a}) \qquad (10.10)$$

$$=B(\underline{a_1})+V(\underline{a_2})$$

如果代理人是资产的所有者，代理人有内在积极性维护和保养资产。为了使代理人选择 $a_1 > 0$，委托人必须根据观测到的利润对代理人进行奖惩，即有 $\beta > 0$。但给定 $\beta > 0$，代理人选择的总努力水平 $a_1 + a_2$ 会大于 $\underline{a}$。这可从以下例子中看出。

例 10.2　设 $B(a_1) = a_1$，$V(a_2) = \sqrt{a_2}$

$C(a_1 + a_2) = (a_1 + a_2 - \underline{a})^2 / 2$，则对任何的 $\beta > 0$，代理人最优化自己的确定性等价收入[6]：

$$V(a_2) + \alpha + \beta_1 a_1 - \frac{1}{2}\rho\beta_1^2\sigma_1^2 - C(a_1 + a_2)$$

一阶条件为：

$$\frac{\partial V}{\partial a_2} - \frac{\partial C}{\partial a_2} = 0$$

$$\beta_1 - \frac{\partial C}{\partial a_1} = 0$$

即

$$\frac{1}{2}a_2^{-\frac{1}{2}} = a_1 + a_2 - \underline{a}$$

$$\beta_1 = a_1 + a_2 - \underline{a}$$

解得 $a_1 = \beta_1 - \left(\frac{1}{2\beta_1}\right)^2 + \underline{a}$，$a_2 = \left(\frac{1}{2\beta_1}\right)^2$

$$a_1 + a_2 = \beta + \underline{a} > \underline{a}$$

一般可令 $a_1^*(\beta_1)$ 和 $a_2^*(\beta)$ 为给定 β_1 时资产所有权属于代理人时的最优努力水平，则总的确定性等价收入为：

$$B(a_1^*) + V(a_2^*) - \frac{1}{2}\rho\left(\beta\sigma_b^2 + \sigma_v^2\right) - C(a_1^* + a_2^* - \underline{a}) \qquad (10.11)$$

若式（10.11）的最大值大于式（10.10）的最大值，则资产所有权就归代理人是最优的产权安排，否则产权归委托人是最优的。

代理人拥有产权有利于激励他在维护和保养资产方面的努力，但增加他的风险成本，包括分享利润的风险成本 $\rho\beta^2\sigma_b^2/2$ 和资产风险成本 $\rho^2\sigma_v^2/2$。最优所有权安排是由二者间的平衡决定的。当 σ_b^2 和 σ_v^2 足够小时，代理人拥有所有权是最优的；当 σ_b^2 和 σ_v^2 太大时，风险成本就主导了激励效应，所有权就应归委托人。在给定 σ_b^2 和 σ_v^2 下，代理人的绝对风险规避度 ρ 愈大，风险成本就愈高，因此，代理人拥有资产所有权的价值就愈低。

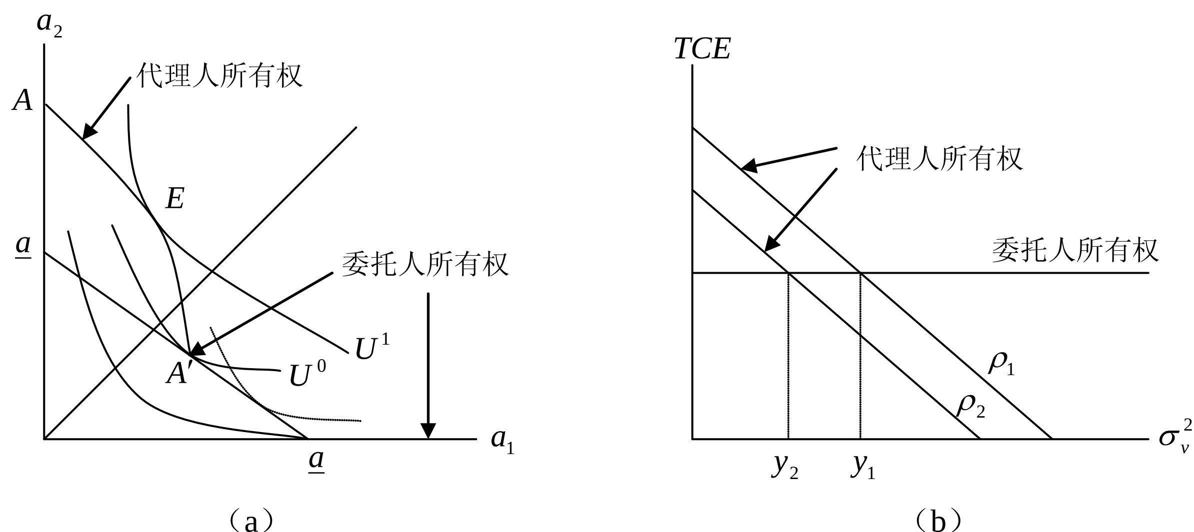


图 10.2　最优资产所有权的安排

图 10.2（a）中，我们用 AA^1 线表达在代理人拥有所有权下的努力供给曲线，委托人选择 β 达到最高的无差异曲线，E 点为均衡点。A' 点是委托人所有权下的最优点，此时 $\beta = 0$；横轴也是委托人所有权下的努力供给曲线，此时 $\beta > 0$。显然，给定 σ_b^2，σ_v^2 和 ρ 下，代理人所有权加上最优激励合同是优于无激励合同的委托人所有权（$\beta = 0$），而后者又优于有激励合同的委托人所有权（$\beta > 0$）。

在图 10.2(b)中，我们给出最优所有权安排与 σ_v^2 和 ρ 的关系，这里假定 $\rho_2 > \rho_1$。当 $\rho = \rho_1$ 时，若 $\sigma_v^2 < y_1$，代理人所有权是最优的；否则，委托人所有权就是最优的。当 $\rho = \rho_2$ 时，若 $\sigma_v^2 < y_2$，代理人所有权是最优的，否则，委托人所有权是最优的。而 $\rho_1 < \rho_2$ 意味着有 $y_1 > y_2$，因为图 10.2(a)中的虚线对应着较高的 σ_b^2，σ_v^2 或 ρ 的无差异曲线，代理人所有权不是最优的。

下面，我们仍用出租车的所有权安排来说明上述理论。如果仅就调动司机多拉客人、多赚钱的角度看，出租车公司应对司机实行激励合同，如承包制。但是，当出租车属于公司所有时，激励合同会使司机不注意汽车的维护保养，甚至会导致司机对车辆的过度使用。如果公司希望司机在维护保养方面花上足够的精力，则承包制就是不可取的。但若司机的收入与赚钱的数量无关，则他就会没有积极性去赚钱了，甚至会把过多的精力花在维护保养上。一个方法是让司机买下汽车，同时又实行承包制，此时司机只上缴固定的管理费。

这种方法会增加司机面临的风险（收入风险和汽车价值变动的风险）。但是，当市场波动不太大时，激励效应将主导风险效应。但若市场波动较大时，如小汽车市场价格

是很不确定的，司机就很可能不愿把车买下。

10.4 第二职业限制制度

公司职员除在公司干活外，还会把其它精力用于所谓第二职业或者称为“外部活动”。如公司的销售人员可能会在销售本企业产品的同时为其它企业销售产品，维修人员可能会在下班后用本企业的设备揽一些私活赚外快；而办公室人员也可能在上班时间打一些私人电话等。我们看到，许多公司或单位都对其职员的第二职业或外部活动作出一些限制性规定，而这些规定又对于不同类型的工作有所不同。如对于那些本职工作有明确的责任规定的代理人如销售人员拥有较多的自由度从事其它的活动，而对于团队型的工作人员的自由度允许又小一些。下面，我们用前面提出的模型对此现象加以解释。

先假定，委托人选择完全禁止代理人从事某项活动比起监督或部分限制该项活动来说要容易。如禁止雇员从事第二职业的规定是常见的，这倒不一定是由于第二职业会损害本职工作，而是因为禁止性制度比起有限度的允许（如允许 10%的时间）从事第二职业在实施上更容易一些。

假设：代理人有有限 N 个可以从事的外部活动，它们都可以为代理人带来一些好处。用 $K=\{1,\cdots,N\}$ 表示所有这些潜在外部活动的集合。用 A 表示委托人允许代理人从事的活动集合，则有 $A\subset K$ 。

代理人可按自己的意愿从事 A 的任一活动，但不能从事 A 之外的活动。假定代理人只有一种本职工作。

令 a 为代理人在本职工作上的努力程度， a_k 为他在第 k 项外部活动上的努力水平（当 $k\notin A$ 时， $a_k=0$ ）， $V_k(a_k)$ 是第 k 种外部活动带给他的确定性收益。

设 $V_k(a_k)$ 严格凹，且 $V_k(0)=0$ ，还设总的成本函数 $C(a+\sum_k a_k)$ 是严格凹的，故代理人的净成本为

$$C(a,a_1,\cdots,a_N)=C(a+\sum_k a_k)-\sum_k V_k(a_k) \tag{10.12}$$

假定：委托人的支付函数以及可观测到的信息分别为：

$$B(a,a_1,\cdots,a_N)=Pa$$

$$x(a,a_1,\cdots,a_N)=a+\varepsilon \tag{10.13}$$

其中 ε 是均值为零，方差为 σ^2 的随机变量。由式（10.13），外部活动并不给委托人带来直接收益，而委托人也观测不到代理人在外部活动上的努力水平。

现在，委托人的问题是选择什么样的佣金 β 和允许什么样的外部活动集 A。下面的分析分两步进行：首先，第一步给定 β 时 A 的最优选择，用 $A(\beta)$ 表示其结果；第二步再给出最优的 β 。

给定 β 和 A，代理人将根据最大化自己的确定性等价收入来选择 $(a,a_1,\cdots,a_N)$ ，其

确定性等价收入为：

$$\alpha+\beta a+\sum_A V_k(a_k)-\frac{1}{2}\rho\beta^2\sigma^2-C(a+\sum_A a_k)$$

设存在内点解，则一阶条件为：

$$\beta=C'(a+\sum_A a_k) \quad （10.14）$$

$$V_k^{'}(a_k)=C'(a+\sum_A a_k)$$

所以有 $\beta=V_k^{'}(a_k)$ （10.15）

这意味着代理人在第 k 种外部活动上的努力水平仅仅依赖于 β，是与 A 无关的（假定有 $k\in A$）。同时，总努力水平 $a+\sum_A a_k$ 也与 A 无关。所以，给定 β，若允许从事更多一些外部活动，代理人会把更多的努力从本职工作转到外部活动。当然，这个结论与（10.13）中的线性假设及完全替代性的成本函数假设有关。

现假设代理人参与约束是等式成立的，则他在外部活动中获得的收益每增加一个单位，委托人支付给代理人的工资就可以减去一个单位，如通过调整固定工资 α，使参与约束仍成立。

委托人的目标是选择 A 去最大化总的确定性等价收入[7]。

第 k 项外部活动的收益为 $V_k[a_k(\beta)]$，其机会成本显然为 Pa_k （由式(10.13)），故最优的 A 为：

$$A(\beta)=\{k\in K|V_k[a_k(\beta)]>Pa_k(\beta)\} \quad （10.16）$$

即当且仅当 $V_k[a_k(\beta)]>Pa_k(\beta)$，委托人才允许代理人从事第 k 项活动。见图 10.3。

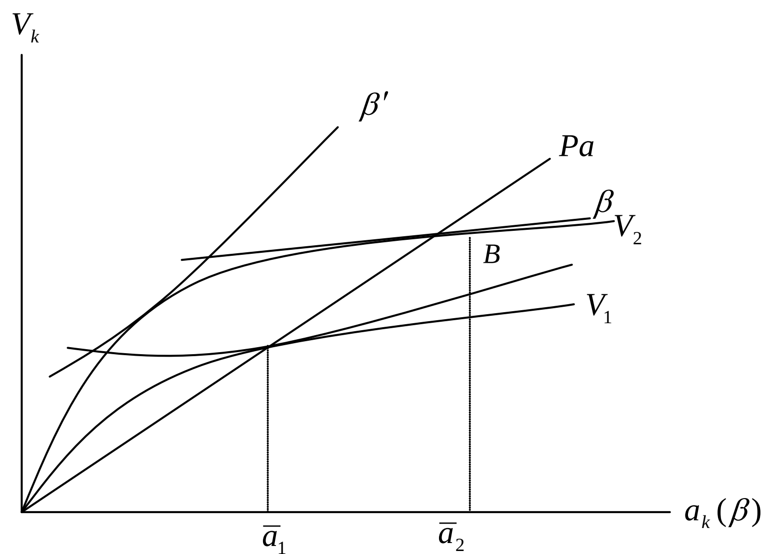


图 10.3 最优 $A(\beta)$ 的决定机理

在图 10.3 中，直线 Pa 是本职工作的收益函数，其斜率为 P。曲线 V_1 和 V_2 分别是两种外部活动的收益函数。由于在区间 $[0,\bar{a}_k]$ 中有 $V_k > Pa$，故两种外部活动在 $a_k \in [0,\bar{a}_k]$ 时都是存在价值的。但在给定 β 下，由式（10.15）代理人会选择 $a_1(\beta)$ 和 $a_2(\beta)$，此时有 $V_1[a_1(\beta)] > Pa_1$，$V_2[a_2(\beta)] < Pa_2$，所以，仅第 1 种活动是值得进行的，第 2 种活动不值得进行下去。

我们从图 10.3 中还可看出，随着 β 的上升，$A(\beta)$ 会扩大，即代理人从本职工作中得到的边际收益愈大的话，代理人就会被允许从事更多的外部活动。在图 10.3 中，β' 比 β 大，此时允许代理人从事第二种外部活动。

之所以有这种规律，是由于 β 是代理人从事外部活动的机会成本。而且给定 V_k 是凹函数，代理人在第 k 项活动上的时间 $a_k(\beta)$ 随 β 的上升而下降，所以，较高的 β 将使有更多的 a_k 满足 $a_k \in [0,\bar{a}_k]$。

我们可以用下述在现实中观察到的制度设计来说明上述理论。我们在实际的制度中可以看到这样的现象：代理人如果在其本职工作中的责任规定愈明确无误，则他去从事外部活动的自由度就愈大。在极端的情形，当 $\beta = P$ 时，委托人就不会去限制代理人从事外部活动了。如当出租车公司实行司机承包制，此时司机对自己的业绩负有全部的责任，公司就不会去限制司机从事其它活动了。这时，司机完全由自己决定何时出车，何时收车，甚至是出车呢还是干私活，相反，对于政府官员来说，他们的 β 接近于零，就不能自由地上下班。譬如，我们看到，对于公司或事业单位的办公室人员，都会规定他们必须按时上下班[8]。

下面来确定最优的 β。委托人问题是解如下最大化问题：

$$\max_{\beta} = \left[Pa + \sum_{A(\beta)} V_k(a_k) - \frac{1}{2}\rho\beta^2\sigma^2 - C(a + \sum_{A(\beta)} a_k) \right]$$

$$s.t. \beta = C'\left(a + \sum_{A(\beta)} a_k \right)$$

$$\beta = V_k^{'}(a_k)$$

故有如下一阶条件：

$$\beta = \frac{P}{1+\rho C''\sigma^2} \tag{10.17}$$

这是式（10.8）的一个特例。我们注意到，最优的 β 与 $V_k(\cdot)$ 是无关的，原因是 $V_k(\cdot)$ 只是对固定工资 α 产生影响。

在式（10.17）中，σ^2 愈大，即本职工作愈难以监督时，对代理人的激励就愈小。结合前面关于 $A(\beta)$ 的结论，在本职工作愈难以监督的情形，对代理人的激励就愈小，故对代理人从事外部活动的限制就愈多。在极端的情形，若委托人能完全观测到代理人

在本职工作上的努力，此时对应于$\sigma^2=0$，则委托人会让代理人对自己的工作负完全责任，$\beta=P$，代理人就可以自由地从事任何外部活动了。相反，若本职工作不可能得到监督，此时$\sigma^2=\infty$，则有$\beta=0$，此时委托人将会禁止代理人从事任何的外部活动。此时，若不禁止，则代理人会将其精力全用于外部活动。

在现实中，我们看到，对于本职工作易于监督的代理人来说，他们会享有较多的外部活动自由，甚至允许从事第二职业。而对于本职工作难以监督的情形，就必须对代理人从事外部活动加以约束。对企业的从事 R&D 的技术人员，他们的工作有较大的不确定性，难以对其实施有效监督。若允许他们从事第二职业，他们就完全可能将在本企业搞出来的成果卖给别的企业，所以企业一般不允许从事 R&D 的人员去从事第二职业。

10.5 专业化分工制度及工种设计

对于前面提出的各任何模型，也可以用于解释专业化分工制度的形成及工种设计的原因。专业化分工指不同的代理人从事不同的工作的安排，而工种设计（grouping tasks into jobs）指将不同的任务划分成不同的工作，从而让不同的代理人从事它们，如将四个任务分成两种工作，让两个不同的代理人从事。专业化分工和工种设计被统称为工作设计（job-design）。

首先对专业化分工制度作出解释。假定有两种工作，以及两个代理人。用$k=1,2$表示两种工作，用i和j表示两个代理人；用a_{1i}和a_{2i}分别表示代理人在第 1 种工作和第 2 种工作上的努力，用a_{1j}和a_{2j}分别表示代理人j在第 1 和第 2 种工作上的努力。进一步假定两个代理人是具有相同的特征的，有关i的结论也适于j。

假定努力的成本函数为$C[a_{1i}+a_{2i}]=\dfrac{(a_{1i}+a_{2i})^2}{2}$，代理人努力的总的期望收益函数为$B(a_{1i}+a_{1j})+B(a_{2i}+a_{2j})$，且$B(\cdot)$是凹函数，可观测信息为$x_1=(a_{1i}+a_{1j})+\varepsilon_i$和$x_2=(a_{2i}+a_{2j})+\varepsilon_i$，即$x_1$反映的是在第 1 种工作上的总努力信息，$x_2$反映在第 2 种工作上的努力。其中$\varepsilon_1$和$\varepsilon_2$是服从独立正态分布，且均值为零，方差各为$\sigma_1^2$和$\sigma_2^2$的随机变量。

我们首先研究没有专业化分工，每一个代理人两种工作都要干的情形。称为共同负责情形[9]。

设代理人i的工资为

$W_i=\alpha_i+x_{1i}x_1+\beta_{2i}x_2$

当我们为了让代理人在两种工作上都有投入，即$a_{1i}>0$，$a_{2i}>0$，就要求在最优化条件中有$\beta_{1i}=\beta_{2i}=\beta_i$，这叫做“等报酬原则”。由于假定两个代理人是相同的，故有$\beta_i=\beta_j=\beta$。于是，我们得到代理人i的确定性等价收入为

$$CE_i = \alpha_i + \beta(a_{1i} + a_{1j} + a_{2i} + a_{2j}) - \frac{1}{2}\rho\beta^2(\sigma_1^2 + \sigma_2^2) - C(a_{1i} + a_{2i})$$

而总的确定性等价收入为

$$TCE_1 = B(a_{1i} + a_{1j}) + B(a_{2i} + a_{1j}) - \rho\beta^2(\sigma_1^2 + \sigma_2^2) - C(a_{1i} + a_{2i}) - C(a_{1j} + a_{2j})$$

委托人选择 B 最大化这个总的确定性等价收入，约束条件是代理人的下述一阶条件：

$$\beta = C'(a_{1i} + a_{2i}) = a_{1i} + a_{2i}$$

$$\beta = C'(a_{1j} + a_{2j}) = a_{1j} + a_{2j}$$

将此约束代入，则委托人的问题就是

$$\max_{\beta} 2[B(\beta) - C(\beta) - \frac{1}{2}\rho\beta^2(\sigma_1^2 + \sigma_2^2)]$$

得一阶条件：$\beta = \dfrac{B'}{1 + \rho(\sigma_1^2 + \sigma_2^2)}$

假定代理人 i 负责工作 1，代理人 j 负责工作 2。由于 ε_1 和 ε_2 是独立的，充足统计量的结论意味着必有 $\beta_{2i} = \beta_{1j} = 0$，即不应用相对业绩作比较（见第 8 章）。总的确定性等价收入为：

$$TCE_2 = B(a_{1i}) - \frac{1}{2}\rho\beta_{1i}^2\sigma_1^2 - C(a_{1i}) + B(a_{2j}) - \frac{1}{2}\rho\beta_{2j}^2\sigma_2^2 - C(a_{2j})$$

代理人的约束条件为：

$$\beta_{1i} = C'(a_{1i}) = a_{1i}$$

$$\beta_{2j} = C'(a_{2j}) = a_{2j}$$

则得到最大化一阶条件：

$$\beta_{1i} = \frac{B'}{1 + \rho\sigma_1^2}$$

$$\beta_{2j} = \frac{B'}{1 + \rho\sigma_2^2}$$

将这两种制度加以比较，我们得到：$\beta < \beta_{1i}$，$\beta < \beta_{2j}$。即专业化分工使对两个代理人的激励都强化了，而且，专业化分工下的总的确定性等价收入是严格大于共同负责下的总的确定性等价收入的[10]。

所以，专业化分工是对共同负责制度的一个帕累托改善。

在专业化分工制度里，每个代理人只面临一种风险(σ^2)；但在共同负责制里，每个代理人都面临两种风险（$\sigma_1^2 + \sigma_2^2$），这样风险的降低反过来通过提高 β 强化代理人更加努力地工作。

现在看一个极端的情形，若委托人没有办法对其中一项工作加以监督，则专业化的好处就更加突出了。如假定第 2 种工作不可加以监督（此时 $\sigma_2^2 = \infty$），则在共同负责制

里，为了使代理人在这项工作中投入精力，最优化条件要求委托人对两个代理人都实行固定工资。即 $\beta = 0$；这样，每个代理人将总的努力水平 $\underline{a}$ 最优分配于两种工作。但是，在专业化条件下，委托人只需对代理人 j 实行固定工资，而对代理人 i 实行激励工资，即 $\beta > 0$；代理人 j 的努力水平为 $\underline{a}$，而代理人 i 的努力水平严格大于 $\underline{a}$。

下面再分析工种设计问题。

假设有四种工作任务 $k = a,b,c,d$ 分配于两个代理人去完成，而每一个代理人干其中两件。设观测任务 a 和 b 的方差都为 σ^2，观测任务 c 和 d 的方差都为 $m\sigma^2$。问题是委托人应如何最优地分配任务。这里来看两种方案，第一方案为：代理人 i 干 a 和 b，代理人 j 干 c 和 d；第二方案为：代理人 i 干 a 和 c，代理人 j 干 b 和 d。

在第一方案中，有 $\beta_i = \dfrac{B'}{1+2\rho\sigma^2}$，$\beta_j = \dfrac{B'}{1+2m\rho\sigma^2}$

总风险成本为

$$\begin{aligned} RC(1) &= \frac{1}{2}\rho\left[\frac{B'}{1+2\rho\sigma^2}\right]^2 2\sigma^2 + \frac{1}{2}\rho\left[\frac{B'}{1+2m\rho\sigma^2}\right]^2 2m\sigma^2 \\ &= \rho\left[\frac{B'}{1+2\rho\sigma^2}\right]^2 \sigma^2 + \rho\left[\frac{B'}{1+2m\rho\sigma^2}\right]^2 m\sigma^2 \end{aligned} \tag{10.18}$$

在第二方案中，有

$$\beta_i = \beta_j = \frac{B'}{1+\rho(1+m)\sigma^2}$$

总风险成本为

$$RC(2) = \rho\left[\frac{B'}{1+\rho(1+m)\sigma^2}\right]^2 (1+m)\sigma^2 \tag{10.19}$$

比较式（10.18）与（10.19），当 $m = 1$，即四种任务观测方差相同，则 $RC(1) = RC(2)$；若 $m \neq 1$，则 $RC(1) < RC(2)$。故若委托人有四种任务分配于两个代理人干，且其中每两种任务的观测方差相同，则最优工种设计是将具有相同方差的任务分配于同一代理人。一般地，工种设计应根据任务的监督的难易来进行的，最易于监督的任务交给一个代理人，而最难以监督的交给另一个代理人。即若

$\sigma_a^2 < \sigma_b^2 < \sigma_c^2 < \sigma_d^2$，那么，$a$ 和 b 就应交给同一个人，而将 c 和 d 交给另一个人[11]。

注释：

[1]这里利用了矩阵$[C_{ij}]$和Σ对称性。

[2]相当于$\sigma_2^2=\infty$，$\sigma_{12}=0$，即测度a_2的误差很大，而从x_1中又得不到有关a_2的丝毫信息。

[3]此时，Milgrom 与 Roberts（1992）将其称为“等报酬原则”。这个结果$\beta_1=\beta_2$还要求有假设$x_1=a_1+\varepsilon_1$和$x_2=a_2+\varepsilon_2$这种线性假定。一般地，当$x_1=u_1(a_1)+\varepsilon_1$，$x_2=u_2(a_2)+\varepsilon_2$时，一阶条件$\beta_1u_1^1(a_1)=C_1$，$\beta_2u_2^1(a_2)=C_2$，当两种努力在成本函数上完全替代时，$C_1=C_2$，则有$\beta_1u_1^1(a_1)=\beta_2u_2^1(a_2)$。

[4]这种现象的解释可以还有其它的理论。

[5]这里$Var(0)=\beta^T\Sigma\beta$，$\beta=\begin{pmatrix}0\\0\end{pmatrix}$，且仍然假定$C(\underline{a})=0$。

[6]与 10.1 中对比，这里代理人的收入中增加资产价值$V(a_2)$。

[7]在第 8 章中，我们就已经知道委托人最大化自己的支付实际上意味着最大化总的确定性等价收入。当然，这与第 8 章中的假定条件是相同的。见第 8 章中的 8.6 节。

[8]$A(\beta)$表面上看起来与代理人的风险规避度ρ，观测变量x的方差σ^2是无关的，但实际上这些变量通过β才与 A 发生作用的。

[9]此时，由于共同负责，可能会出现团队生产中的“搭便车”。这个问题可以由后面提出的通过给予委托人剩余索取权的方式打破平衡约束（breaking budget Constraint）来避免。

[10]使用包络定理可直接比较两个确定性等价收入可获得这一结论。

[11]这就是说，此时应“精耕细作”而不是“广种薄收”，与其弱化对所有人的激励还不如只弱化对一部分人的激励。假定一部分任务可监督，另一部分不可监督。若每一个人都既干可监督的任务也干不可监督的任务，则不会有人有激励去努力工作。相反，若将不可监督的任务交给一些人去干，另可以监督的任务交给另一些人干，则此时只是对前一些人没有激励而已。

第十一章　监督与效率工资制度

委托人可以通过投入一定的资源对代理人的行为进行监督，因而在一个完整的委托—代理理论中，还应对委托人通过投入资源来增加他对代理人行为观察的信息的可能作出假设。委托人既可以通过花费资源去监督代理人，也可以让渡一定的产出去激励代理人。只有在边际上花费的资源价值小于这种监督带来的信息价值时，边际上增加资源去完成对代理人的监督才是值得的。相反，当监督边际成本大于监督边际收益时，就应通过让渡一定的产出去激励代理人。这正是我们在第 7 章的开始就提到过的思想。在这一章里，我们把监督引入委托—代理问题，从而形成一种对第 7 章的一种补充。我们在两个领域中来讨论这一问题，一个是第 7 章中就已介绍的效率工资模型，另一个是前面第 8 章以及随后几章中提出的激励理论范围内。

在效率工资领域中，我们用委托人发现代理人“偷懒”的概率作为委托人的选择变量（战略），而在激励理论的框架里，委托人的选择变量（战略）是关于代理人行为的观测误差。

11.1 效率工资模型中的最优监督问题

在第 7 章里，我们曾经介绍了效率工资。在那里，我们是以介绍不完全信息动态博弈模型的角度去说明效率工资，但在本章，与第 7 章不同的是，我们将从制度设计的角度重新对效率工资展开讨论。经济学家最早是在发展经济学中提出效率工资（efficiency wage）的概念的。新古典经济学的传统教义认为工人的工资率取决于其边际生产率，但经济学家们在发展中国家看到的经常是边际生产率取决于工资率。即不是老板按照工人的边际生产率来给工人定工资，而是工人根据老板开出的工资率水平去提供边际生产率。有人认为，在发展中国家，由于工人们的营养状况普遍较差，以及知识和技能的掌握程度较低，老板给予工人们较高的工资率水平可能会改善工人的营养水平，提高工人的知识和技能水准（工人们在较宽裕的生活环境中可以去加强学习），从而提高边际生产率。实际上，这种现象不仅在发展中国家可以看到，“边际生产率取决于工资率”在发达国家也存在，这是后来人们进一步的发现。经济学家们还发现，企业给工人支付的工资率实际上还高于工人的保留工资率，这种现象如何解释呢？

一些经济学家提出，高工资率可能是企业在不能观察到工人可能的偷懒行为时对工人努力工作的一种激励，即当企业不可能完全对工人的行为加以监督时，高工资率就成为偷懒工人一旦被发现后被解雇的高机会成本；这样，工资率愈高，这种机会成本就愈大，工人偷懒的动机就愈小（Solow, 1979；Shapiro 与 Stiglitz，1984）。这就是效率工资理论的基本思想。

我们先来考察一种静态模型，设代理人是风险中性的，其效用函数为

$U(w,a)=w-C(a)$

其中 w 是工资率，a 是代理人（工人）的努力水平，$C(a)$ 是工人的（努力）成本函数，假定 $C'>0$，$C''>0$，$C(0)=0$。

假定 a 只取 $a=0$（偷懒）和 $a=1$（努力工作）两个数值。

假定：企业只能通过直接发现工人偷懒的行为来确认工人的努力水平偏低，不能通过其它方式做到这一点（譬如，不能通过总产出或成本的高低来确认某一工人是否努力工作，因为总产出或成本是所有工人行为的结果，难以与某一个工人的行为直接对应）。

按照前一章的结果，企业此时不可能对工人实行激励工资，只能实行固定工资。设 P 是在工人偷懒时，企业能发现工人偷懒的概率。当工人不偷懒时，工资率为 w，努力成本为 $C(1)$，效用为 $U(w,1)=w-C(1)$。当工人偷懒的时候，一旦被发现会被立即开除，效用就是保留效用 w_0，即 $U(w_0,0)=w_0$；若偷懒又不被发现，效用为 $U(w,0)=w$。故工人选择偷懒的时候，期望效用为 $Pw_0+(1-P)w$。

当且仅当如下条件成立的时候，工人选择努力工作：

$w-C(1)\geq Pw_0+(1-P)w$

得 $w(P)\geq w_0+\dfrac{C(1)}{P}$ （11.1）

如果企业能完全监督工人，则参与约束是等式成立的，有 $w=w_0+C(1)$。于是，式（11.1）指出在企业不能对工人行为进行完全的监督时（$P<1$），为使工人努力工作，企业就要给工人开出大于保留工资率与努力成本之和的工资率水平。当监督愈困难时，P 愈小，企业支付的工资率就愈高。当 $P=0$，完全不能监督，则任何工资率都不能使工人努力工作，令 $\Delta P=w+C(1)/P-[w+C(1)]=C(1)(1-P)/P$，则 ΔP 就是在不能完全监督时，企业为诱使工人努力工作而在工资中加入的一种“贿赂金”。

下面，我们来考察 P 的决定。当委托人投入资源加强监督时，P 就愈大。如企业花更多的钱雇用监工和安装电子监视系统，则 P 就会提高。我们以下假定发现偷懒的概率 P 对应的监视成本为 $M(P)$，且 $M'(P)>0$，$M''(P)>0$，$M(0)=0$，$M(1)=\infty$[1]。

这样，代理成本就包含有“贿赂金” ΔP 和监督成本 $M(P)$ 两个部分。如果提高 P，则 ΔP 将下降，但会增大监督成本。于是，最优的 P 应使总代理成本 $AC(P)$ 最小化，其中

$$AC(P)=C(1)\frac{1-P}{P}+M(P)$$

一阶条件为：

$$-\frac{1}{P^2}C(1)+M'(P)=0$$

在图 11.1 中，我们给出了一阶条件的几何图像。

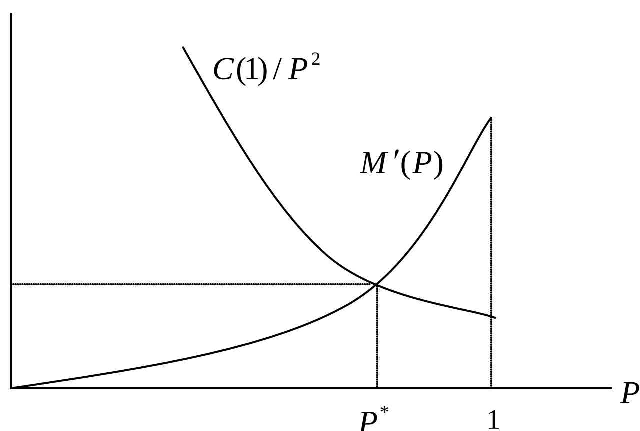


图 11.1　最优监督

在图 11.1 中，$C(1)/P^2$ 是 P 增大时的边际贿赂金节约，即边际收益，而 $M'(P)$ 是 P 提高时的边际监督成本。式（11.1）意味着在最优监督处提高 P 的边际收益等于边际成本。P^* 是最优监督概率。显然，当 $C(1)$ 上升时，P^* 也会上升[2]。

以上分析实际上还潜在假定了企业从雇用工人中得到的总收益 $\beta(1)$ 是大于工资支出 $w(P^*)$ 的，否则企业就不会雇用工人。

下面，我们用该模型来解释"高薪养廉"。设 a 是代理人（如公务员）的"廉洁努力水平"，$a=1$ 表示廉洁，$a=0$ 表示腐败（滥用职权）。用 w 表示公务员的本职工作工资水平，而他被开除公职后的市场工资率为 w_0，滥用职权带来的额外收入为 $R(\lambda)$。设 P 是滥用职权被发现的概率，一旦被发现滥用职权，他还要被罚 F 数量的罚款。其中 λ 是公务员的权力度量。假设有 $R'(\lambda)>0$，即权力愈大，滥用职权就获得愈大的额外收入。于是，公务员滥用职权带来的期望收益就为：

$$P(W_0-F)+(1-P)[W+R(\lambda)]$$

所以，公务员选择"廉洁"的条件是

$$W \ge P(W_0-F)+(1-P)[W+R(\lambda)]$$

得：$W \ge (W_0-F)+\dfrac{1-P}{P}R(\lambda)$

给定 W_0、F 和 $P(P<1)$，公务员的权力愈大，则要使他选择"廉洁"就应给予他愈高的工资率 W[3]。

11.2 激励理论中的最优监督问题

在上面一节，我们假定不能观察到工人努力水平本身（只能"发现"工人偷懒），故在应实行的固定工资下研究了最优监督问题。在这一节中，我们在激励理论中讨论最优监督问题。此外，可以在一定程度上观察到代理人的努力水平，但这种观察是存在误

差的，如果委托人投入更多的资源，则可以将这种观测误差缩小。

假定代理人的努力水平为a时，委托人的期望收益为$B(a)$，观测信息记为$x = a + \varepsilon$，ε是服从正态分布，均值为零，方差为σ^2的随机变量。设工资合约是$w = \alpha + \beta x$。当代理人具有不变绝对风险规避度时，总的确定性等价收入为：

$$TCE = B(a) - \frac{1}{2}\rho\beta^2\sigma^2 - C(a)$$

代理人的激励相容约束为

$$\beta = C'(a)$$

委托人的最优化一阶条件为

$$\beta = \frac{B'}{1 + \rho C''\sigma^2}$$

故σ^2愈大，代理人得到的激励就愈小，而总的确定性等价收入就愈低（或代理成本就愈高）。

下面设$B(a) = a$，$C = a^2/2$，则$B' = 1$，$C' = a$，$C'' = 1$。

故总的确定性等价收入为：

$$TCE = \beta - \frac{1}{2}\rho\beta^2\sigma^2 - \frac{1}{2}\beta^2 = \frac{1}{2}\left(\frac{1}{1 + \rho\sigma^2}\right)$$

如果委托人可以降低σ^2，则总的确定性等价收入就能提高了。如何才能降低σ^2呢？一种方法是加强监督，但这是要花费资源成本的，故委托人需要在收益与成本之间进行权衡。设$M(\sigma^2)$是监督的成本函数，$M(0) = \infty$，$M(\infty) = 0$，$M'(\sigma^2) < 0$，$M''(\sigma^2) > 0$，则净的收益（福利）为

$$W(\sigma^2) = \frac{1}{2}\left(\frac{1}{1 + \rho\sigma^2}\right) - M(\sigma^2)$$

委托人通过选择σ^2最大化该净收益。一阶条件为：

$$\frac{1}{2}\frac{\rho}{(1 + \rho\sigma^2)^2} = -M'(\sigma^2) \tag{11.2}$$

式（11.2）的左边是降低σ^2的边际收益，右边是对应的边际成本。

当$\rho = 0$时，没有代理成本，故委托人不必去降低σ^2。但若$\rho > 0$，则通过降低σ^2就可以强化对代理人的激励，即提高β，从而降低代理成本。

我们在图 11.2 中，降低σ^2的边际成本随σ^2的下降而上升，降低σ^2的边际收益也随着σ^2的下降而上升，但边际成本的上升速度快于边际收益的上升速度（当$\sigma^2 \to 0$时，边际收益趋向于$\rho/2$，但$-M' \to \infty$）；若$\rho = \rho_1$，最优的σ^2是y_1；当$\sigma^2 > y_1$时，边际收益就大于边际成本；而当$\sigma^2 < y_1$时，边际成本大于边际收益。类似地，若$\rho = \rho_2$，则最优的σ^2为y_2。若边际成本曲线处处都低于边际收益曲线，委托人就会选择$\sigma^2 = 0$，另外，若边际收益曲线处处低于边际成本曲线，则委托人就不会花费任何成本去降低

σ^2。

代理人的绝对风险规避度ρ与最优方差的关系不是单调的。

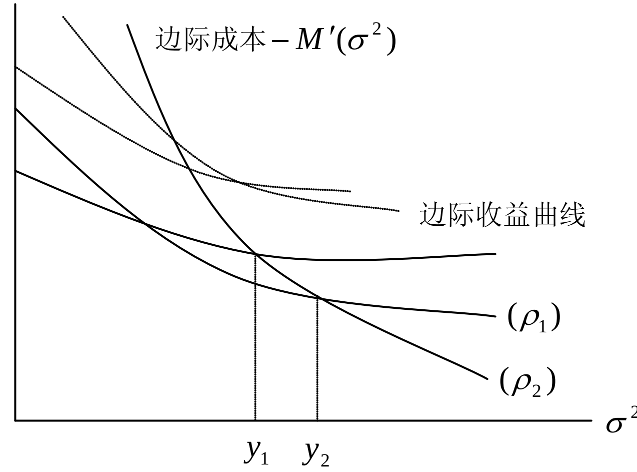


图 11.2　最优化方差

当$M=1/\sigma^2$，$-M'=1/(\sigma^2)^2$，则最优方差为$\sigma^2=\frac{\sqrt{2}}{\left[\sqrt{\rho}-\sqrt{2}\rho\right]}$，若$\rho<1/8$，则$\frac{\partial\sigma^2}{\partial\rho}<0$；若$\rho>1/8$，则$\frac{\partial\sigma^2}{\partial\rho}>0$。这是因为，$\rho$与$\sigma^2$通过两个方面共同对 TCE 产生作用，一是通过对最优激励参数β的影响，二是通过对风险成本的直接影响。虽然ρ对 TCE 的总的效应是负的，但由式（11.2），当ρ足够小时，ρ对降低σ^2的边际收益的效应是正的，ρ的上升降低最优的σ^2（如图 11.2 中的两条虚线表示的边际收益曲线所表达的那样）；当ρ较大时，ρ对降低σ^2的边际收益的效应是负的，ρ的上升提高最优的σ^2（如图中的两条实线表示的边际收益曲线所示）。

特别地，在下述两种情况委托人不必去监督。第一种情况是$\rho=0$（即代理人是风险中性的），此时代理人对自己的行为负完全的责任（$\beta=1$），因而有内在的动力努力工作，第二种情况是$\rho=\infty$，即代理人是完全风险规避的，代理人拿固定的工资（$\beta=0$），此时监督是没有意义的（当然，给定委托人是风险中性的）。

下面，我们再来看一些决定σ^2的因素是如何影响上述结论的。假定有$B(a)=Pa$，$C(a)=ba^2/2$及$M(a)=\frac{m}{\sigma^2}$。在这种情形，可将P解释为努力的边际生产率，而b是努力的边际成本系数（b愈大，边际成本就愈高），m是刻画监督难度的系数。

此时，由最优化条件有$\beta=ba$，即$W^*=\beta/b$。委托人选择β的最优化条件意味着：

$$\beta = \frac{B'}{1+\rho C''\sigma^2}$$

$$= \frac{P}{1+\rho b\sigma^2}$$

总的确定性等价收入为：

$$TCE = Pa - \frac{1}{2}\rho\beta^2\sigma^2 - C(a)$$

$$= \frac{P^2}{2b}\left(\frac{1}{1+\rho b\sigma^2}\right)$$

委托人的最优化一阶条件为：

$$\frac{P^2}{2}\frac{1}{(1+\rho b\sigma^2)^2} = \frac{m}{(\sigma^2)^2}$$

解得：$\sigma^2 = \left[\sqrt{\frac{P^2\rho}{2m}} - \rho b\right]^{-1}$ （11.3）

求导得：

$$\frac{\partial\sigma^2}{\partial P} < 0，\quad \frac{\partial\sigma^2}{\partial m} > 0，\quad \frac{\partial\sigma^2}{\partial b} > 0$$

这说明代理人的边际生产率愈高，监督带来的边际收益愈高，则委托人就愈有积极性去监督；代理人努力边际成本愈高，任何给定激励下的努力供给就愈低$(a = \beta / b)$，并且给定σ^2下最优的激励就愈低$[\beta = P(1+\rho b\sigma^2)^{-1}]$，监督的边际收益就愈低，委托人监督的积极性自然也愈低；监督愈困难，监督的边际成本愈高，委托人去监督的积极性就愈低了[4]。

注释：

[1]事实上不需要假设 $M(0)=0$ 和 $M(1)=\infty$，只需假设不投入资源进行监视时 P 就充分小，而当投入资源充分大时有 $P=1$。

[2] $C(1)$ 上升时，给定 P 下的贿赂余额也上升，故监督的边际收益就愈大。

[3]但是，公务员的保留工资 W_0 也可能与权力 λ 呈正相关，因为权力愈大的公务员，就会有愈大的关系网，当他离开政府后，其就业机会就愈好，就愈可能获得较高的市场工资。

[4]在第三点上，较多的监督指委托人选择较低的 σ^2，不是指委托人在监督上花较多的支出或精力。

第 12 章　新制度经济学中的企业理论

新制度经济学认为企业是一种团队生产（team product），而“团队”（team）是指企业中的工人是一组代理人，他们独立地选择努力水平，但却生产出一个总的产出，每个代理人对产出的边际贡献是依赖于其他代理人的努力的，是不可单独观测的[1]。

新制度经济学的一个代表人物 Holmstrom 提出的团队生产理论框架是下面分析的出发点。

在团队生产中，由于个人努力不可观测，将会出现个人的“搭便车”偷懒行为。由此，应引入一个监督者（monitor），同时为了对监督者本身要有激励，则监督者就应该是剩余索取人（residual claimant）。这就解释了古典企业的存在原因。但 Holmstrom（1982）曾证明，对于团队生产中的偷懒问题，可通过激励机制来解决。他认为，监督者实际上是委托人，他的作用并不是监督团队成员，而是通过剩余索取权打破了预算平衡（breaking budget）。

下面，我们来看看这个结论是如何得到的。

设有 n 个代理人 $i=1,2,\cdots,n$；代理人 i 选择行动 $a_i \in A_i=(0,\infty)$，个人成本为 $C_i(a_i)$，$C_i^{'}(a_i)>0$，$C_i^{''}(a_i)>0$，且 $C_i(0)=0$。

记 $a=(a_1,\cdots,a_n)=(a_i,a_{-i})$ 为行动向量，其中 $a_{-i}=(a_1,\cdots,a_{i-1},a_{i+1},\cdots,a_n)$。生产函数为 $x=x(a)$，且 $x(a)$ 是严格递增的可微凹函数，$x(0)=0$。

制度安排 x 在 n 个代理人之间分配，记 $S_i(x)$ 为代理人 i 的所得份额。设代理人是风险中性的，故效用函数为 $u_i(S_i,a_i)=S_i(x)-C_i(a_i)$ [2]。

现在的问题是：是否存在一种制度安排，使在纳什均衡下的分配是一个帕累托最优的分配。为此，我们先看在预算平衡时的情形，此时对任意产出 x 有

$$\sum_{i=1}^{n} S_i(x)=x \qquad (12.1)$$

在式（12.1）两边对 x 求导：

$$\sum_{i=1}^{n} S_i^{'}(x)=1 \qquad (12.2)$$

代理人 i 将选择 a_i 最大化其效用 $U_i=S_i(x)-C_i(a_i)$

一阶条件为：

$$S_i^{'}(x)x_i^{'}=C_i^{'}(a_i),\quad i=1,\cdots,n \qquad (12.3)$$

但是，帕累托最优的行动向量 a^* 应满足：

$$a^*=a\,rg\max_{a}\left[x(a)-\sum_{i=1}^{n} C_i(a_i)\right]$$

一阶条件为

$x_i^{'} = C_i^{'}(a_i)$， $i = 1, \cdots, n$ （12.4）

由式（12.3）和（12.4），如果纳什均衡又是帕累托最优的，则必有 $S_i^{'} = 1$（$i = 1, \cdots, n$），但这与预算平衡约束式（12.2）矛盾。故有如下结论：

满足预算平衡约束式（12.2）的纳什均衡努力水平是严格小于帕累托最优努力水平的。即在预算平衡下不可能有帕累托最优均衡。这是因为，每个代理人只能得到自己的边际产出的一部分（$S_i^{'} < 1$ 份额），故代理人都没有积极性选择最优的努力水平使边际产出等于边际成本。这就是搭便车问题（free-rider problem）。

下面，我们在这里通过引入一个委托人来解决搭便车问题。这就是古典资本主义企业取代合伙制企业。当把预算约束改为：

$$\sum_{i=1}^{n} S_i(x) \le x \quad （12.5）$$

再看如下分配方式：

$$S_i(x) = \begin{cases} b_i, & 若 x \ge x(a^*) \\ 0, & 若 x < x(a^*) \end{cases} \quad （12.6）$$

在这里，a^* 由条件（12.4）决定，是帕累托最优努力向量。也就是说，据此方案，若总产出大于或等于帕累托最优产出，代理人 i 得到 b_i；若总产出小于帕累托最优产出，则代理人一无所得。

不难证明在这个分配方式中，纳什均衡是帕累托最优的。在帕累托最优处，$x(a) - \sum_{i=1}^{n} C_i(a_i)$ 达到最大化。一般该剩余会大于零[3]。

于是，有 $x(a^*) > \sum_{i=1}^{n} C_i(a_i^*)$，这样，就可以找到一些 b_i。

满足：$b_i > C_i(a_i^*)$ 且 $x(a^*) = \sum_{i=1}^{n} b_i$， $i = 1, \cdots, n$。

给定其他代理人选择 a_{-i}^*，若代理人 i 选 $a_i < a_i^*$，则有

$$x(a_i, a_{-i}^*) < x(a^*, a_{-i}^*)$$

由式（12.6）知，$S_i(x(a_i, a_{-i}^*)) = 0$，$u(0, a_i) = -C_i(a_i) \le 0$，但是，若他选 $a_i > a^*$，则 $x \ge x(a^*)$， $S_i(x) = b_i$， $u(b_i, a_i^*) = b_i - C_i(a_i) \le b_i - C_i(a_i^*) > 0$。

显然，他的最优选择是 a_i^*，这样，a^* 就是一个纳什均衡。

除此之外，我们还可以作出其它的一些帕累托最优均衡方案。

譬如，如果所有代理人都是相同的，制度安排要求每一个代理人在开始就缴纳一个保证金，数额为 $A_i = \frac{(n-1)}{n} x(a^*)$，然后按 $S_i(x) = x$ 分配。这也是打破了预算平衡的，因为有

$$\sum_{i=1}^{n} S_i x) = nx > x$$

这一制度下有 $S_i^{'} = 1$，然后据式（12.3）和（12.4），在均衡下是帕累托最优的。每个代理人得到的实际净收入就为[4]

$$x(a^*) - A_i = \frac{x(a^*)}{n}$$

Arrow（1985）曾指出，上述打破预算平衡的制度可能有多重均衡，因更多的行动组合都是最优反应。他指出："如果有人偷懒，为了达到目标产出，其他人就会更加努力地工作（大于帕累托最优水平）是值得的。因此，这个方案并不能保证最优结果，尽管它允许最优结果。"

但是，在现实中也能找到类似的制度，如企业给个人支付固定工资加上团体奖金，奖金只有在一个给定的目标（类似这里的 $x^* = x(a^*)$）达到后才能获得。如果企业的业绩不佳，在极端情形董事会全部被解职[5]。

但是，如果事后实际产出小于 $x(a^*)$，此时将 x 抛掉并不符合团队利益。但是，如果代理人事前就预期事后不会抛掉小于 $x(a^*)$ 的 x，从而"团体惩罚"不可实施的话，就会出现搭便车。这说明，这个制度不是子博弈精炼均衡，不满足动态一致性。现在，我们通过引入一个有剩余索取权的委托人。如果 x 小于 $x(a^*)$，则产出全归委托人所有。譬如，在上面那个制度中，委托人收取保证金，然后按 $S_i = x$ 支付给代理人，委托人的剩余为 $\left[x(a^*) \frac{(n-1)}{n} \right] n + x - nx > 0$。并且，委托人本身不是团队成员。

所以，从这个理论角度看，委托人的真正作用不是监督代理人，而是保证不满足预算平衡的团体激励方案的实施。古典式资本主义企业与合伙制企业相比较的优点就在于它可以使用一些在合伙制下不可行的激励方案。

在确定性的环境中，均衡下没有剩余留给委托人。

譬如，在第一种制度下有 $\sum_{i=1}^{n} b_i = x(a^*)$，而在第二种制度中有 $\frac{x(a^*)}{n} \cdot n = x(a^*)$。现在的问题是，如果委托人事前就预期自己将一无所得，就没有人愿去当委托人。但 Holmstrom 证明在不确定的环境里，这类团体激励是很有效的。我们下面来看他的思路。

假定产出还受到外生因素 θ 的影响，故有 $x = x(a, \theta)$。使用分布函数的参数化模型（见第 8 章），把 x 当作随机变量。

令 $F(x,a)$ 和 $f(x,a)$ 分别为 x 的分布函数和分布密度，假定 $F(x,a)$ 满足一阶随机占优条件和凸性条件（见第 8 章）。设计如下制度

$$S_i(x) = \begin{cases} S_i x, & 若 x \geq \bar{x} \\ S_i x - k_i, & 若 x < \bar{x} \end{cases} \tag{12.5}$$

其中$k_i > 0$，$\sum S_i = 1$。显然有$\sum_{i=1}^{n} S_i(x) \le x$。这个制度规定：如果产出达不到目标$\bar{x}$，则每个代理人受到$k_i$数额的处罚，否则，总产出全部分配给代理人。设代理人是风险中性的，则代理人的期望效用函数为

$Eu_i = W_i + ES_i x(a) - k_i F(\bar{x}, a) - C_i(a_i)$，$i = 1, \cdots, n$

其中W_i是代理人的初始财富求导。设a^*为帕累托最优努力向量，则

$$a^* \in \arg\max\left[Ex(a) - \sum_{i=1}^{n} C_i(a_i)\right]$$

得$E_i x(a^*) = C_i^{'}(a_i)$，$i = 1, \cdots, n$

这里$E_i x(a) = \dfrac{\partial Ex(a)}{\partial a_i}$，故纳什均衡的充要条件是

$$S_i E_i x(a) - k_i F_i(\bar{x}, a) - C_i^{'}(a_i) = 0,\quad i = 1, \cdots, n \tag{12.6}$$

这里$F_i x(a) = \dfrac{\partial F(\bar{x}, a)}{\partial a_i}$。设分布函数是有界的，且代理人的初始财富水平足够高，则通过调整$\bar{x}$和k_i就可以保证纳什均衡（12.6）的解是帕累托最优的。下面以一个例子对此加以说明。

例 12.1　设$n = 2$，$S_i = \dfrac{1}{2}(n = 1,2)$，$x$是在$[0, 2(a_1 + a_2)]$上的均匀分布，代理人的努力水平决定这个分布区间的上界位置，则

$$f(x, a) = \frac{1}{2(a_1 + a_2)}$$

$$F(x, a) = \int_0^x f(x, a) = \frac{x}{2(a_1 + a_2)}$$

$$Ex(a) = (a_1 + a_2)$$

设成本函数为$C_i(a_i) = \dfrac{a_i^2}{2}$，则帕累托最优努力水平$a_i^*$满足式

$a_i^* = 1$，$i = 1,2$

记$\bar{x} = a_1^* + a_2^* = 2$，则$F(\bar{x}, a) = \dfrac{1}{(a_1 + a_2)}$，纳什均衡条件为：

$$\frac{1}{2} \times 1 + \frac{k_i}{(a_1 + a_2)^2} - a_i = 0,\quad i = 1,2$$

由对称性，均衡时有$a_1 = a_2$，故上式为

$2a_i^2(2a_i - 1) - k_i = 0$，$i = 1,2$

均衡下$a_i = 1$，故得$k_i = 2$，$i = 1,2$。

于是，在产出x小于$\bar{x} = 2$时，将对代理人罚款$k_i = 2$，此时每个代理人的收入为$\dfrac{x}{2} - 2 = (x - 4)/2 < 0$；这个分配方案就可实现帕累托最优。一般情形下，当$k_i$和$\bar{x}$满足等式$k_i = 4\sqrt{x}$，这个分配方案就可以实现帕累托最优（这个结论可留给读者自行去验证，

请注意，当惩罚的临界点$\bar{x}$愈低时，惩罚的数额k_i就愈高。）

如果代理人的初始财富水平不是足够大，而是很有限，则帕累托最优就有可能达不到[6]。在这个例子中，当代理人的初始财富水平为零时，帕累托最优是不可能实现的。在一种极端情形，当产出分布随着n的增加而快速分散的，即对所有x有$\lim_{n\to\infty}|F_i|=0$，则均衡努力水平趋于零。所以，在不确定环境里，代理人的财富约束可能会限制团队的有效规模及实现帕累托最优的可能性。但是，当委托人的初始财富足够大时，通过将“团体惩罚”改为“团体激励”，帕累托最优仍可达到。譬如，这样一个分配方案要求当$x>\bar{x}$时，委托人支付代理人奖金b_i；当$x\le\bar{x}$时，委托人支付代理人$S_i x$，$\sum S_i=1$。则通过适当调整b_i和$\bar{x}$，纳什均衡就可作为帕累托最优出现。

由此，Holmstrom 认为委托人的监督只有在团队规模很大，代理人及委托人都面临初始财富约束及代理人是风险规避的时候才是重要的[7]。

通过直接对代理人行为加以监督，委托人可收集更多信息，对代理人的奖惩不再仅基于团队产出作出，这样，一方面减少了团体惩罚和激励的必要，从而放松了财富约束，另一方面又降低了代理人承担的风险，如采用相对业绩作比较等。

有人不仅考虑了团队生产中的道德风险，而且还考虑了其中的逆选择（能力不可观测），如 Mcafee 与 McMillan（1991）。他们证明：在一定条件下，最优工资合约是团队产出的线性函数；即$S_i(x)=\alpha_i+S_i x$。若不存在逆选择，则帕累托最优就可通过规定$S_i=1$和代理人向委托人缴纳保证金$-\alpha_i$达到[8]。

但若存在逆选择，则最优的S_i小于 1，从而$-\alpha_i$小于没有逆选择时的水平。这是由于通过将代理人的边际工资规定得低于边际产出，委托人可以从高能力的代理人那里榨取信息租金。他们证明，不论委托人是只观测团队产出x，还是观测每个人的贡献，则均衡结果是一样的，即使观测个人贡献没有成本。个人贡献的不可观测并不一定会带来搭便车问题，而监督也并不是消除偷懒的必要手段。这是 Holmstrom 早已指出的（Holmstrom，1982）。但是，他们认为，监督的作用是约束委托人自己而非代理人。这是因为：若有n个相同的代理人，不存在逆选择，则据建立总产出上的最优合约，每个代理人在事前向委托人缴纳$-\alpha=x(a^*)(n-1)/n$的保证金，委托人在事后向每个代理人支付x，因此对n个代理人的总支付为nx，大于产出x本身；产出每增加一单位，委托人要支出n单位；产出愈小，委托人支出就愈少。故委托人可能故意破坏生产使代理人只能达到较低的产量。委托人也可能在生产结束前就携带代理人缴纳的保证金跑掉。解决委托人的这种道德风险的一个办法是：让委托人监督代理人而非收取代理人的保证金。由于在监督情况下，代理人的产出愈高，委托人的剩余就愈多，委托人不会破坏生产，也没有保证金可拿着逃跑（或只有较少的保证金可带走）。

Itoh（1991）在文献中的贡献也值得一提。他的思路与文献中的一般看法不同，他不是将团队生产视为外生的，也不是强调团队生产中的搭便车行为，而是提出这样一个

问题：在什么样的情况下，委托人应通过激励机制诱使代理人互相帮助？这样，团队生产便是一种内生结果了。设有两个代理人，有两种工作，每个代理人主要负责其中一种工作，且两种工作的业绩是不相关的，因此不存在相对业绩的比较问题。委托人要考虑的是：是否应诱使每个代理人除了在自己的工作上努力工作外，还应花一定的精力帮助他的同伴？这个问题可进一步分成两个问题。第一是将一个代理人的收入适当（正相关地）依赖于另一个代理人的业绩是否能够诱使代理人互相帮助？第二是若互相帮助能出现，这样做是否值得委托人去做？对于第一个问题，Itoh 证明，若代理人帮助别人的努力边际成本在帮助努力等于零时为零，如成本函数为$C(a,b)=C_1(a)+C_2(b)$，其中a是自己工作上的努力，b是帮助努力，则答案是肯定的，因为适当帮助别人并不增加自己的边际成本。他证明，这种情形在两种工作的差异较大并且代理人对不同工作有不同偏好时就可能出现。但若两种工作很类似，使得代理人只注重总努力水平，即$C(a,b)=C(a+b)$，则答案就是否定的了。此时，帮助的边际成本在$b=0$时严格大于零，即使对别人很小的帮助都增加自己工作的边际成本，代理人就不会帮助别人。对于第二个问题，他证明，给定第一个问题的答案是肯定的情形，团队工作或相互帮助是最优的一个充分条件就是自身努力和相互帮助在如下意义上是战略性互补的：即来自别人的帮助愈多，自己工作愈努力。在博弈论中，正斜率的反应函数意味着战略互补，负斜率的反应函数是战略替代。战略互补指局中人的战略变动方向相同，战略替代指反向的战略变动。譬如，价格竞争就是战略互补的，你提高价格，我也提价；但产量竞争就是战略替代的，如你提高产量，我就减少产量。战略互补在以下情形更容易出现；即当别人帮助自己干一些较不重要的事情时，自己可集中精力干一些更为重要的事情。于是，我们有如下结论：若两种努力在成本函数上是独立的但在工作上是互补的时，通过激励机制诱使“团队工作”是最优的。即适当帮助努力并不增加多少成本，即可通过使别人更努力工作提高产出水平。进而，即使代理人对来自别人帮助的最优反应是减少自己的努力（即搭便车，是一种“战略替代”），若所导致的自己努力的下降可大大降低适当努力水平的效用成本，团队工作仍会是最优的。如果本职工作十分单调乏味，则这种情况就会出现。正如行为科学发现的那样，丰富多彩的工作内容本身可以激励个人更加努力工作。但是，当第一个问题的答案是否定的时，情况就较为复杂。互补性不再是团队工作最优的充分条件了。委托人必须考虑诱使代理人从事多项工作的成本，这种成本在$b=0$和$b>0$之间是非连续性的。当两种努力在工作上的互补性很强时，最优的选择就可能是两个极端，或者每个人只是专业化自己的工作，或是有很高程度的团队工作。诱使严格专业化的激励机制是每一个人的工资只依赖于自己的工作业绩，诱使高度团队工作的激励机制是每个人的工资主要依赖于团队产出。这一点可以解释现实中观测到的一些企业人事制度。

总之，代理人之间的战略依存方式（互补或替代）和他们对不同工作的态度是决定团队是否最优的两个主要因素。

注释：

[1]此时，相对业绩比较或锦标制度都不适用。

[2]这里潜在假定代理人的初始财富为零。

[3]否则，$a=0$ 就是帕累托最优，我们排除这一无意义的情形。

[4]n 个代理人的事后分配是平衡预算的，因为每个人的净收入之和为 $x(a^*)$。这种制度实际上等价于激励工资 $S_i=\alpha+\beta x$，而其中 $\alpha=-A_i$，$B=1$。

[5]在一个动态模型中，我们可以将式（12.6）作为中断合作的威胁。

[6]除非是 x 的分布十分集中。

[7]以上分析假定代理人是风险中性的。

[8]前面已经知道这一点。